Competition and Market Strategies in the Swiss Fixed Telephony Market. An estimation of Swisscom's dynamic residual demand curve.

Roberto Balmer

University of Rome I

10. December 2013

Competition and Market Strategies in the Swiss Fixed Telephony Market

An estimation of Swisscom's dynamic residual demand curve

Roberto E. Balmer

University of Rome I, Department of Public Economics

JEL-codes: D43, L1, L41, L51, L9

Abstract:

Fixed telephony has long been a fundamentally important market for European telecommunications operators. The liberalisation and the introduction of regulation in the end of the 1990s, however, allowed new entrants to compete with incumbents at the retail level. A rapid price decline and a decline in revenues followed. Increased retail competition consequently led a number of national regulators to deregulate this market. In 2013, however, many European countries (including Switzerland) continued to have partially binding retail price regulation. More than a decade after liberalisation and the introduction of wholesale and retail price regulation, sufficient data is available to empirically measure the success of regulation and assess its continued necessity. This paper develops a market model based on a generalised version of the traditional "dominant firm – competitive fringe" model allowing the incumbent also more competitive conduct than that of a dominant firm. A system of simultaneous equations is developed and direct estimation of the incumbent's residual demand function is performed by instrumenting the market price by incumbent-specific cost shifting variables as well as other variables. Unlike earlier papers that assess market power in this market, this paper also adjusts the market model to ensure a sufficient level of cointegration and avoid spurious regression results. This necessitates introducing intertemporal effects. While the incumbent's conduct cannot be directly estimated using this framework, the concrete estimates show that residual demand is inelastic (long run price elasticity of residual demand of -0.12). Such a level of elasticity is, however, only compatible with a profit maximising incumbent in the case of largely competitive conduct (conduct parameter below 0.12 and therefore close to zero). It is therefore found that the Swiss incumbent acted rather competitively in the fixed telephony retail market in the period under review (2004-2012) and that (partial) retail price caps in place can no longer be justified on the basis of a lack of competition.

1. Introduction

Fixed telephony was of fundamental importance to European telecommunications operators at the beginning of liberalisation at the end of the 1990s. In Switzerland, the incumbent Swisscom's fixed retail telephony revenues (access and traffic) made up for over 70% of its retail revenues in 1999[1]. In a context of overall declining telecoms revenues and fixed telephony prices, this share dropped to below 30% in 2011. Over this period, other telecommunications services such as broadband and mobile have increased in both relative and absolute importance. However, while a number of national regulators started to deregulate retail telephony markets after judging them competitive, in 2013, many European countries, including Switzerland, had partially binding retail price regulation for fixed calls still in place. More than a decade after the liberalisation and the introduction of wholesale and retail price regulation, sufficient data is available to empirically measure the success of regulation overall in recent years and to assess the continued necessity of retail price regulation.

This paper develops a market model based on a generalised version of the traditional "dominant firm – competitive fringe" model. Competitive fringe firms are assumed to be price takers and the dominant firm (Swisscom) acts as price leader able to move first, perfectly anticipating its competitors' supply reactions, to set an optimal unique market price for fixed telephony. The assumption of a "dominant firm" is subsequently relaxed considering also competitive conduct closer to price taking behaviour by introducing a *conduct parameter*.

A simple system of simultaneous equations is developed, allowing direct estimation of the incumbent's residual demand function by instrumenting the market price by incumbent-specific cost shifting variables, such as its number of staff (to control for possible remaining inefficiencies from times when the incumbent was fully state controlled), as well as other variables. Earlier papers have sometimes used indirect methods to estimate residual demand elasticity for a lack of good incumbent cost shifters (see Kahai, Kaserman and Mayo (1996)). The concrete model developed is, however, only valid under strong assumptions, including fixed termination rates close to zero, identical marginal network access costs for all competitors and independence from mobile telephony. Nevertheless, it is shown that these assumptions are reasonable in this context and the period under review (2004-2012).

Unlike earlier papers trying to assess market power in this market, this paper adjusts the market model to ensure a sufficient level of cointegration and to avoid spurious regression results. This necessitates introducing intertemporal effects. Any change in a variable has therefore an immediate same period effect on residual demand, and also a series of effects on future demand. While conduct cannot be directly estimated from this general framework, the concrete estimates show that residual demand is inelastic (long run price elasticity of -0.12). Such a level of elasticity is, however, only compatible with a profit maximising incumbent in case of largely competitive conduct (conduct parameter below 0.12 and therefore close to zero). It is therefore found that Swisscom acted rather competitively in the fixed telephony retail market in the period under review. If the problem of an uncompetitive retail market ever existed, it seems that the entry of alternative operators such as cable and the introduction of regulated wholesale access (carrier pre-selection as well as local loop unbundling) have successfully removed it. This implies that (partially) regulated retail price caps in place in Switzerland can no longer be justified on the basis of a lack of competition, and calls into question whether there are legitimate reasons for their continuation.

---

[1] Swisscom's total fixed voice access and traffic revenues in 1999 amounted to 6.6bn CHF, compared to 1.8bn in mobile services and 0.9bn CHF in broadband. In 2011, fixed voice access and traffic revenues amounted to 1.8bn CHF, compared to 3.4bn CHF in mobile services and 0.6bn CHF in broadband (Swisscom annual reports 1999 and 2011).

Chapter 2 describes the fixed telephony market characteristics and the main assumptions of the model developed in this paper. In addition, it describes the concrete situation in Europe and in particular Switzerland, in terms of market structure and regulation. Chapter 3 outlines the basic analytical framework used to estimate fixed voice traffic (residual) demand. Chapter 4 provides an overview of the input data and Chapter 5 estimates the model. Chapter 6 interprets the results and Chapter 7 provides concluding remarks.

2. Market characteristics and model assumptions

This chapter describes the fundamental characteristics of telecommunications markets which need to be taken into account before building a market model for fixed telephony in Switzerland in subsequent chapters. In particular, the main properties of wholesale markets, where telephony operators buy access products from their competitors to originate and terminate calls, of cost functions and of demand are briefly analysed in light of the relevant literature. In addition, regulation and market structure over the relevant period is reviewed. From these characteristics, the main assumptions of the market model are derived.

Armstrong (2002) describes the fundamental properties of competition between retail fixed telephony operators (see also Laffont and Tirole (2001) and Vogelsang (2003)). Most importantly, operators are not identical. Competitors without infrastructure typically seek access to infrastructure owned by the incumbent to provide services to end-customers (one-way access). As the European incumbents usually are also present at the retail market, they are vertically integrated, while (most) competitors are only active on the retail market[2]. It is shown in this chapter that this situation is inefficient and can lead to foreclosure, which is why access is usually subject to some form of regulation. In addition, operators need to access each others' networks to terminate calls on competing networks (two-way access). While competitive interactions are more complicated, competition problems are also likely to arise, which is why termination is usually also subject to regulation. Wholesale regulation is therefore found to be one of the key drivers of the retail market for fixed telephony. Finally, it is shown that mobile calls do not represent good substitutes for fixed calls and that the fixed arms of firms with their own mobile networks take decisions independently in this context.

One-way access and vertical integration
Telecommunications operators need to have own physical infrastructure (local access network) or access to such an infrastructure to provide fixed telephone calls at users' homes. In most countries today, there are only few fixed access infrastructures allowing for independent fixed call origination (e.g. copper, cable, fibre) while at the time of liberalisation only one such infrastructure was broadly available (copper[3]). Operators without their own infrastructure may ask infrastructure-based operators for access[4]. In this case, so-called "one-way" access is requested, i.e. where the incumbent firm has control over an important input needed by its rivals to compete at retail level, but where it itself needs nothing from other firms. Competitors then pay an "access" charge to the incumbent. In other terms, Vogelsang (2003) relates access charges to cases where networks operate at different hierarchical levels and only one network uses the other to originate (and terminate) calls on the calling party's side. "Interconnection

[2] The case of vertical separation is also analysed in Laffont and Tirole (2001). Typically the risk of exclusion does not exist (while the problem of monopoly power may persist).
[3] Cable networks were able to provide voice calls only later. In Switzerland, main cable operator Cablecom launched its telephone service in June 2004.
[4] It should be noted that wholesale access is usually not sold by cable operators.

charges", instead, are related to two networks at the same hierarchical level, linked in order to enable calls across different networks (providing for example call termination for each other).
Returning to one-way access, the connection used to connect to the telephone network can technically be ensured by two different traditional wholesale access products (Armstrong (2002)). One is "call origination" where the incumbent provides for a telephone connection, and origination of traffic is bought by the competitor. In this case, the local telephone access continues to be operated by the incumbent and only long distance calls (i.e. over different regional points of interconnection where national calls are bundled) would require network elements operated by the access-seeking firms. A second possibility for entrants without their own access network is to provide telephony services to end-users through unbundling of the local loop, where the competitors operate their own telephone service platform up to the end-user[5].

Vogelsang (2003) reviews the relevant literature on one-way access and shows that under vertical integration, an incumbent may not be willing to grant one-way access, i.e. let competitors use the essential facility at reasonable terms, and that a foreclosure problem may exist[6]. Regulators should then impose access prices. Optimal access prices maximising welfare under the constraint that the incumbent breaks even are called Ramsey prices, and are typically above marginal costs as in this case the competitors have to make a fixed cost contribution to the incumbent. Given the complexity of implementing Ramsey pricing (e.g. lack of information on elasticities and competitive reactions), such pricing has not yet broadly been implemented by regulators. Instead, telecom regulators have usually used cost-based access price determination mechanisms (in particular long run incremental cost (LRIC) where the investment costs for the different services are shared with rivals on cost basis). The level and impact on the market of a regulated fixed call origination charge is analysed in the next section.

Call origination costs

The telephony operators' total as well as marginal costs can be divided in two broad categories. One is the network costs faced by a firm to establish the necessary telephone connections and handle the traffic and the other is costs such as customer care and administration. For alternative operators without their own network, the network access costs are represented mainly by the (potentially above marginal cost) regulated access prices. The incumbent a priori faces only its marginal cost curve. With price independent demand this may be irrelevant for the market outcome as the incumbents' perceived marginal cost corresponds to its opportunity cost on the wholesale market. With price dependent demand[7] (and when each competitor has a price dependent hinterland of loyal customers unaccessible to the other operator), instead, this can lead to inefficiencies. In this case, when an incumbent is vertically integrated and retail competitors rely on its wholesale products, Inderst and Peitz (2012) show that the incumbent charges lower uniform retail prices in equilibrium than its competitors and that it has a higher market share (partial foreclosure)[8]. This would be the case whenever regulated access prices are above marginal network cost[9]. Conversely, Gans and King (2005) show that when access prices for non integrated competitors are set at marginal network cost (and integration does not provide strategic benefits or costs) this implies competitive neutrality. When the regulated access price is set at marginal

[5] In Switzerland markets referring to both mentioned access products were found to be not effectively competitive, respectively the incumbent has never denied to have significant market power, and regulation is actively imposed (LRIC prices).
[6] With differentiated goods this problem may partially persist when demand is price dependent (see Inderst and Peitz (2012)). In fact, the incumbent would set an access price above marginal cost and would set lower retail prices than the competitor in order to attract more customers in the loyal customer segment where it faces no opportunity cost but only its marginal network cost.
[7] Which can be assumed to be the case
[8] As the incumbent in its hinterland incurs only its marginal network cost and not an (above marginal cost) opportunity costs corresponding to the other operators' wholesale cost, an asymmetry between the firms arises.
[9] It should be noted that there can also be a so-called "softening effect". When upstream access is provided a customer lost on the retail market may then be recovered partially via the upstream market. This may make integrated firms less aggressive on the retail market than other firms.

network costs, the competitor can therefore compete at exactly the same grounds as the incumbent and there is a priori no reason for inefficient allocation even in case of product differentiation and price dependent demand. When there are two firms, the retail market outcome after investment is then equivalent to a competitive model of a market without vertical integration.
The relevant one-way access prices in Switzerland are regulated at cost-based LRIC prices (i.e. above marginal network cost of the incumbent). Access prices are required to be non-discriminatory, meaning that the incumbent's retail arm faces a marginal network cost equal to LRIC prices - as all its access seeking competitors. A drop in the cost and price of the regulated product would therefore have the effect to reduce perceived marginal costs in an identical way for both the incumbent and the competitors. The costs taken into account by the cost model of the regulator for "regional origination" of calls are the elements to construct a telephone network and originate calls (not considering the construction of the physical access network which is a separate increment). Essentially, these costs are related to the operation of a local and regional telephone switching platform and corresponding backhaul lines. It can be assumed that, differently to the physical access network, a significant part of these costs are operating and maintenance costs and that such costs may depend on scale. In this particular case, the regulated long run incremental costs are therefore assumed to be close to marginal costs (of an efficient operator). In the model it is therefore assumed for simplicity that LRIC prices for this wholesale product correspond largely to marginal network cost of origination and that access seekers therefore compete on the same grounds as the access providers when considering call origination. The aspect of vertical integration of the incumbent can therefore be ignored.

Competitors with their own infrastructure have not been considered in this analysis. In Switzerland, cable operators started to offer their own telephony services in 2004. In addition, competitors started to provide telephony services based on local loop unbundling in 2008. However, it can reasonably be assumed that all these operators are similarly efficient in operating telephone networks as an efficient incumbent[10], and that they therefore face the same marginal costs for call origination and have the same marginal network cost drivers. Such costs are in this paper assumed to be largely exogenous and to be given by the regulated access price for call origination. Additional network costs to complete calls (conveyance and termination on different networks) as well as other costs related to the operation of a telephony networks (core network maintenance, customer care, administration, etc.), however, also need to be considered by competitors.

Marginal network costs are therefore considered identical for origination for all operators and exogenous. All other costs are supposed to be able to differ across operators (e.g. long distance network elements, customer care, etc.). For overall marginal costs, a positive slope is assumed in the segment under consideration. In other words, as in most markets, it is assumed that at some point diseconomies of scale set in and marginal costs increase, essentially due to limited capacity. When some threshold of traffic is reached, operators may have to install more backhaul capacity which may be very costly and lead to very high marginal costs for the level of demand which could trigger such an investment. Similar considerations can be made for other costs. For example, the main Swiss fixed telephony competitor Sunrise seemed to face an important increase of costs at some point. The increase in demand has made it necessary for the firm to recruit about 170 new call center operators (on 830 existing[11]) in only three months and to buy additional (detached) office spaces. For these reasons it is therefore plausible that for the operation of a telephone network alone, overall marginal costs of the operators may be increasing.

It can therefore be concluded that the assumptions that the hypotheses' that:

---

[10] The regulators costs are calculated for an efficient incumbent, therefore considering always the most efficient technology to provide the service.
[11] Source: http://www.20min.ch/digital/news/story/Bei-Sunrise-sitzt-der-Chef-im-Call-Center-27828675

i) regulated LRIC origination prices correspond to an efficient operator's marginal network cost (and all infrastructure-based operators are assumed to be efficient)
ii) overall marginal cost is increasing with output

are plausible in the context under analysis implying that operators compete on the same grounds in the retail market and that competitive firms would increase their output with the market price.

Two-way access

In order to provide a successful phone call from a users' home, an operator not only needs access to the customer, but must also ensure conveyance of its calls and delivery by the competitor providing services to the user called (termination). Under two-way access, of which termination is an example, operators use each others' bottleneck inputs to be able successfully terminate calls and compete with each other. The need for regulation is here less straightforward than under one-way access. Armstrong (2002), Briglauer (2010), Valletti and Calzada (2005) and Valletti and Genakos (2011) propose market models taking into account termination costs and revenues. The literature review of Vogelsang (2003) shows that two issues may arise: collusion (especially under symmetric interconnection relationships) and exclusion (especially when relationships are strongly asymmetric). While every operator may have exclusive control of calls terminated on its subscriber base it must at the same time reach an agreement with competitors on a termination rate in order for its customers to be able to call the competitors' subscriber base. The equilibrium outcome largely depends on the operators' positions and market parameters. As both collusion as well as exclusion undermine competition, this situation has led in Europe to price regulation of termination rates (usually cost-based).

In Switzerland, termination on the incumbent's fixed network is similarly actively price regulated (LRIC) since (nearly) the beginning of liberalisation. While there is no direct regulation of the other fixed operators' termination fees, the incumbents' regulated fixed termination standard offer for fixed operators contains a reciprocity clause[12]. Such clause means that bilateral termination rates of Swisscom with its fixed competitors are identical and - as the incumbent also offers regulated transit to fixed termination of third operators[13] - that all other fixed termination rates are indirectly regulated by arbitrage at or near the incumbents' rate. It is therefore assumed that all operators receive the same regulated fixed termination fee. It should be noted here that from a practical point of view only operators with independent telephone networks receive termination fees. Therefore, for telephone operators without an own network as via unbundling or cable, the termination fees are received by the network operating firm (the incumbent).

No active regulation is, however, in place for mobile termination in Switzerland. This is due to the litigation based regulatory system, where a market intervention is only be triggered upon request by an operator. There has never been an upheld lawsuit of an operator against another operator to ask for regulated access to termination on its mobile network yet. Therefore such termination rates are still unregulated. Unsurprisingly, Swiss mobile termination rates were (among) the highest in Europe in the period under consideration. For example in January 2010, Switzerland had the highest average mobile termination rate, 10.7€cent per minute, compared to a European average of 6.3€cent[14]. For comparison, the price of the regulated fixed termination charge in Switzerland was then significantly below 1€cent per minute[15]. Mobile termination charges were in absolute terms therefore more than ten times higher than fixed termination charges. This particular regulatory framework may imply that the cost to terminate mobile calls was of far greater importance to fixed operators' profits than any possible

[12] Source: http://www.swisscom.ch/en/wholesale/products/voice-services.html and http://www.swisscom.ch/dam/swisscom/de/ws/documents/D_IC-Dokumente/D_IC_LB_Terminierungsdienste_V1-0.pdf
[13] Swisscom Transit Termination
[14] Source: BEREC, BoR(10)30rev1
[15] Source: Swisscom Price Manual 7-2

(fixed) termination revenue[16], especially when considering a homogenous good of voice traffic towards all national networks. Therefore, when considering an overall market for national calls, a strongly simplified framework - necessary as will be shown for estimation reasons - could for the above reasons foresee full abstraction from fixed termination revenues and costs of fixed operators, while taking into account mobile termination rates on the cost side.

It can therefore be concluded that the assumptions that:

i) all Swiss fixed operators face the same fixed regulated termination charge
ii) the fixed termination charge is zero and can be ignored in the model in terms of revenues and costs
iii) the mobile termination charges are high and taken into account as costs for fixed operators

are plausible, implying that all fixed operators compete on the same grounds in the retail market.

Horizontal differentiation

Unlike other telecommunications services, telephone calls from the fixed network usually have no particular features that allow for horizontal differentiation, and vertical differentiation seems largely excluded as well[17]. The differentiation characteristics considered in Kahai, Kaserman and Mayo (1996) referred to the so-called “carrier selection” model, where a customer wishing to use an alternative operator had to dial additional digits before the telephone number (this could be interpreted as negative quality). However, with the introduction of “equal access” wholesale offers in the U.S. (“carrier pre-selection”), no additional digits were necessary anymore for the preselected carrier. By 1993, in the U.S. 97% of wholesale offers were converted to equal access. In Switzerland in the period under review, carrier preselection could be opted for by competitors, foreseeing that customers would not have to dial additional digits. Such source of differentiation is therefore absent in this case and it can be assumed that fixed telephony is a largely homogeneous good. In the case of a homogeneous good, a single equal retail price for telephone services can be assumed as otherwise rational consumers would tend to switch. This assumption may be limited by the fact that switching may be complicated and lengthy or by differentiation of brands. Only in recent years - and therefore unlikely to affect the present analysis of the years 2004-2012 - a further source of differentiation may be present with bundles (in particular where some operators can offer additional services such as TV and others not). In addition, on the market there are a number of price plans and components and comparability and transparency may not be a given (e.g. two-part tariffs, on- and offnet tariffs, day and night/weekend tariffs, bucket subscriptions, etc.). Nevertheless, in light of the above it seems reasonable to assume for simplicity that fixed telephony services are a homogeneous good. This paper abstracts therefore from residual differentiation possibilities.

Fixed-mobile substitution

When designing a market model to estimate market power in the fixed telephony market possible substitutes need to be taken into account. Fixed telephony has one major potential competitor, which is mobile telephony. Typically, mobile access and calls are more expensive than fixed network calls, as they provide the feature of mobility. Vogelsang (2010) reviews the literature on fixed-mobile substitution and concludes that mobile telephony is a substitute for fixed telephony (positive cross-price elasticity) in wealthy countries[18] (calls only, while the situation is less clear for access). Moreover, the author argues

[16] This can also be seen, for example from the case of Sweden, where the regulatory authority has published industry fixed voice service revenues in the market (including fixed charges, fixed to fixed and fixed to mobile calls, international calls and other) as well as fixed termination revenues. Fixed termination revenues corresponded to less than 5% of the service revenues in 2012. Source: PTS: “The Swedish telecommunications market 2012”
[17] Commercialised HD calls and video calls are still rare.
[18] Where the fixed network coverage is not more extended.

that with increasing mobile penetration and decreasing prices, substitutability should further increase over time. While theoretical work is often inconclusive, empirical work, most prominently Briglauer, Schwarz and Zulehner (2010) hints towards substitution. The authors analyse monthly telephony market data in Austria from 2002 to 2007. They find that for residential users, there is a positive cross-price elasticity between fixed and mobile for national calls. They consequently argue therefore that fixed and mobile calls should be considered part of the same market. However, the Austrian market is notoriously very competitive and results do not have to translate to Switzerland under the period of review. Most importantly, the Austrian regulator was to date the only European regulator finding a joint retail market for fixed and mobile voice telephony (Austrian regulatory authority case AT-2009-0881). All other European regulators have to date not come to this conclusion in their national markets. A survey by BEREC (2012) shows that the main reason for these regulators not to define a common retail market was the existence of different product characteristics between fixed and mobile offers, in particular different price levels and the mobility of mobile services. BEREC (2012) cite an Analysys Mason study[19] estimating that in Western Europe fixed calls are cheaper by 37% than equivalent mobile calls. In addition, the conclusion of Briglauer, Schwarz and Zulehner (2010) does not apply to access. In this case, it is argued that probably for quality differences and the possibility to share costs among household members there seems to be no fixed mobile-substitution.

Given the above, a model of the fixed telephony market (calls) should probably also take into account supply and demand for mobile calls. When considering a long period of time retrospectively and in countries with particularly high mobile prices or low mobile penetration and, this may, however, not be necessary.

Regarding prices, a study by the Finnish Communications Regulatory Authority (2009) finds (at about mid-period of the dataset which will be considered) that from 19 European countries, Switzerland had in by far the highest mobile telephony prices with a medium basket expenditure of around 70€ per month compared to a European average of 42€ per month. Conversely, fixed telephony prices were closer to the average[20]. As in addition a long retrospective period is considered (2004-2012), the present model can reasonably assume the absence of fixed-mobile substitution. A factor supporting this hypothesis is that the duration of fixed calls in Switzerland was found to be more than double the duration of mobile calls before 2010, indicating a different use of the two technologies[21].

Independence of fixed telephony

Even in a framework where absence of fixed-mobile substitution is assumed, it is, however, not clear whether the fixed and mobile markets are fully independent. Mobile operators have revenues from mobile termination fees paid by fixed operators. When an integrated operator offers mobile and fixed network services it might, when setting a fixed retail price, internalise effects of mobile termination revenues. For instance, lower fixed voice prices and higher volumes in the market could increase an operators mobile termination revenues. In practice, in the case of Swisscom, its high mobile termination rate had implied that its fixed division payed consistent termination fees to its mobile division[22]. As both fixed operators with mobile networks had separate business units for fixed and mobile networks with separate objectives for a large part of the period under analysis, it is assumed in this paper for simplicity that fixed operators operate independently even when they have mobile arms. They are therefore

[19] Mobile-only households: fixed voice will all but disappear in some Central and Eastern European countries – September 2010

[20] OECD Communications outlook 2011. E.g. 140 fixed calls, VAT included, Switzerland: 26$, OECD average: 27$

[21] For example, in its 2010 decision on mobile termination the Swiss federal court of Justice (B-2050/2007) argued that the fixed call duration in the period under review was more than double the mobile call duration. This seemed to clearly indicate a different use of the two technologies, the more expensive, mobile solution on the go and the cheaper stationary solution at fixed locations.

[22] The Swiss federal court states that in 2005 Swisscom (Fixed division) had paid transfer payments to the three mobile operators (including Swisscom mobile) of more than 100m CHF. See RPW 2010/2.

assumed to pay mobile termination rates as any other operator and to set prices to maximise their business unit profit. In addition, it can be assumed in this case that fixed operators perceive mobile termination rates as (largely) exogenously given, as for example in Briglauer, Götz and Schwarz (2010).

Retail price structure

In this section it may be convenient to review market prices before reviewing the theory on different tariff structures. As in most countries, standalone telephony services for residential customers in Switzerland includes telephony access and a corresponding access fee, as well as usually standard call set-up and per minute charges. Telephone access has in older studies been shown to be price inelastic. For a review of the literature on fixed voice access own price elasticities see Ward and Woroch (2010) and Gassner (1998). It can therefore be assumed that competition works mainly via the usage-based part of the tariff. This assumption seems to be confirmed in the Swiss market where in 2013 a standalone telephone access costs about 25 CHF per month at all major fixed telephone operators (Swisscom, Sunrise, Cablecom). This retail access fee corresponds to the maximum charge for universal service (including VAT) as defined in the Swiss Telecommunications Act[23] and in particular the Ordinance on Telecommunications Services[24] (article 22). Volume based tariffs differ, however. For example, Swisscom charges 0.08 CHF/min for calls to the fixed network (50% of it during the night and on weekends) and 0.35 CHF/min for calls to the mobile network (0.30 CHF/min during the night and on weekends). Swisscom's per minute charge again corresponds to the maximum charge for universal service. Cablecom, instead, foresees no charges for calls to the fixed network while it charges 0.40 CHF to all mobile networks. Sunrise charges 0.06 CHF/min to the fixed network (free during the night and on weekends) and 0.35 CHF to mobile networks (0.30 during the night and on weekends). In addition, these calls incur call setup fees. In particular, Cablecom and Sunrise charge a setup fee of 0.12 CHF per call, while Swisscom does not charge such a fee. In summary, all major operators charge the same fixed fee, but differentiate their price plans according to call set up fee, day and night/weekend tariffs, fixed and mobile calling prices, and a free initial number of minutes. In addition, a number of subscriptions are available.

The simplest possible model setup regardind the retail tariff structure would be a linear tariff. This would mean the absence of a fixed fee (two-part tariff) and the presence of a single usage-based charge per minute (i.e. taking into account any call setup charges, day and night/weekend charges, all fixed and mobile calling prices and free minutes). As can be seen in the next chapter the number of telephone accesses (even excluding Internet based telephone services such as Skype) is by far exceeding the number of households in Switzerland (about 5m accesses on average during the period under review against about 3.2m households). It can therefore be assumed that - even though there may be a number of business lines - in a large majority of households an active telephone access was present. In addition, demand for access is typically inelastic. It is therefore reasonable to think that competition for traffic largely independent from access.

Nevertheless, a two-part tariff nature of pricing in fixed telephony markets may have indirect effects. The presence of a fixed fee may correspond, in fact, to some level of transfer (based on bargaining power) between consumers and firms, while the traffic based prices aim to maximise rent extraction (see for example Inderst and Peitz (2012)). Of relevance for this particular case, Growitsch, Marcus and Wernick (2010), have shown that the fixed part of the tariff has a strong negative correlation with the level of the (fixed) termination rate. To see this, the literature on "waterbed" effects has to be considered. Termination brings not only costs for operators but also revenues. The revenues a consumer brings to an operator from other users calling him and terminating on its operators' network may be relevant. When such revenues are important competition may imply that an operator (e.g. a mobile operator with

---

[23] http://www.admin.ch/opc/en/classified-compilation/19970160/201007010000/784.10.pdf
[24] http://www.admin.ch/opc/en/classified-compilation/20063267/201212280000/784.101.1.pdf

high termination charges) may lower its retail prices to attract consumers in order to have access to the related termination revenues. Some authors argue that this “waterbed” effect mostly affects the fixed fee, which is reduced in such cases. Conversely, when termination rates are lowered by regulatory authorities, termination revenues per user are more limited and the possibility of granting discounts to attract customers on this basis are reduced. Overall, it is unclear whether cost or revenue effects will dominate. Some literature has developed to understand the extent of the waterbed effect. For instance, Genakos and Valletti (2011) estimate that lower mobile termination rates would lead to higher overall mobile retail prices. Most authors find, however, that lower termination rates lead to a decline in retail prices but to a lower extent (implying a partial or incomplete waterbed effect). This does not mean that all components of the retail price would decrease. Most authors, like Growitsch, Marcus and Wernick (2010), assume that the fixed fee would be increased in case of a reduction of termination rates (negative correlation). However, as in this model also the absence of (relevant) fixed termination charges and fixed termination revenues is assumed, it can be assumed that there are no such transfers over the period under consideration. This means that fixed fee revenues and costs related to fixed voice access can be largely ignored. Regarding the usage based prices as in Armstrong’s basic model (1998) linear retail tariffs (i.e. per minute prices) are considered.

To construct a comparable linear usage-based charge a homogeneous composite good of national voice traffic with a single per minute price for calls in all national networks is constructed. In reality, tariffs are as shown varying according to different parameters (e.g. day/night, call set-up fees, bucket plans, fixed/mobile network termination, etc). As fixed termination rates are assumed to be the same for all operators and to be (near) zero it is assumed that there are no cost-based reasons for discrimination of prices between different fixed networks. Given the large number of (other) dimensions of usage-based retail prices, any empirical model needs to make some form of tariff aggregation in order to ensure comparability. A possible way to calculate a single average per minute national tariff for fixed telephony is to divide all usage based national call revenues by the number of national traffic minutes (average revenue per minute (ARPM)). Doing so aggregates all dimensions explained earlier.

It should be noted that while this paper concentrates on traffic one aspect of access is taken into account which is that the number of accesses increases the extent of the network and potential users called. In this sense only (and exogeneously), it is expected that accesses influence traffic.

Given the above, it can therefore be concluded that the following assumptions are reasonable:

i) fixed voice access fees are not considered
ii) a single national per minute retail fixed voice calling price is considered

Regulation

In this chapter the regulation of fixed telephony operators in Europe and in particular in Switzerland is briefly reviewed. In the European Union, there is a clear trend towards deregulation of retail telephony markets as the European Commission stopped considering this market in need of ex-ante regulation in 2007[25]. Most member states have consequently started to withdraw regulatory remedies in this market in the years following the recommendation. According to Cullen, in 2013, six out of twenty-seven EU member states still regulated their telephony retail markets (Belgium, Bulgaria, Cyprus, Hungary, Poland, Portugal). Most regulations include some form of price control. Some national regulatory authorities (NRAs) have also demonstrated recently that competitive problems in this market persist, most recently Bulgaria in 2013[26]. The EU Commission has accepted such analyses’ in several cases

[25] Commission recommendation on relevant product and service markets within the electronic communications sector susceptible to ex ante regulation - C(2007) 5406
[26] Case BG/2013/1421

but advised the national regulatory authorities to reassess the situation in the next round of market analysis, as wholesale remedies may become sufficient to ensure retail competition in this market (e.g. case BG/2013/1421). In Switzerland, the regulatory framework has never foreseen the possibility of traditional asymmetric regulation of operators in the retail market. As has been described earlier, however, national retail price caps are in force for the universal service operator (incumbent) and seem to be binding (both for telephony access as well as fixed voice calls) for a standalone standard offer. There are, however, a number of subscriptions offered which imply that the prices charged by Swisscom may be substantially lower. Overall, the question of competition in retail markets for telephony services (national calls) seems in any case to be controversial across Europe.
Unlike the retail market, wholesale markets have been subject to active regulation not only in the rest of Europe but also in Switzerland. In particular, LRIC cost-based access prices for fixed call origination (since 1998) and for unbundled access to the local loop (only since 2008) are set by a regulator. However, regulation applies, unlike in the EU, only if parties cannot agree on terms of contract and there is a formal complaint to the Swiss Communications Commission (litigation based regime for regulation). Decisions of the authority may then be appealed at the Federal Administrative Court. Only after such a final decision regulated wholesale prices become binding which may make wholesale regulation in Switzerland slow in the sense that the market impact may be effective only years after the period under review. For example the fixed voice origination charges for the years 2000-2003 were lowered by 30% only in late 2006 upon regulatory intervention.

Market structure
In order to choose an appropriate market model the broad structure of the fixed voice telephony market needs to be analysed. It has to be pointed out that on top of the physical access line (copper, cable, FTTH), different technologies can be used to provide phone calls. While this may include traditional (PSTN and ISDN) technologies, this may also include proprietary Voice over IP (VoIP) solutions (typically used by competitors over cable or DSL) as well as Voice over broadband (VoBB), where calls are made using a technology via the IP layer and “over the top” of a retail broadband connection (e.g. Skype) with a platform operated over the Internet. The latter technology is typically unable to guarantee quality of service. It is assumed in this paper that VoBB cannot offer sufficient quality of service to be considered a valid substitute. This choice is also necessary as no reliable data is available.

At the industry level, it can be seen that the number of active fixed voice accesses in Switzerland is slowly declining (Figure 1). This is not true for all types of accesses. While traditional PSTN/ISDN based telephone accesses are slowly declining both for the incumbent as well as for retailers (buying corresponding wholesale origination solutions from the incumbent), there is a steady increase of the number of proprietary VoIP accesses based on DSL or Cable. This increase is related to the market entry of Cable operators in 2004 as well as to the unbundling activities of competitors (in case of unbundling a fully independent VoIP telephony platform can be operated). This overall industry decline might be compensated to some extent by VoBB fixed telephone accesses offered by operators such as Skype.

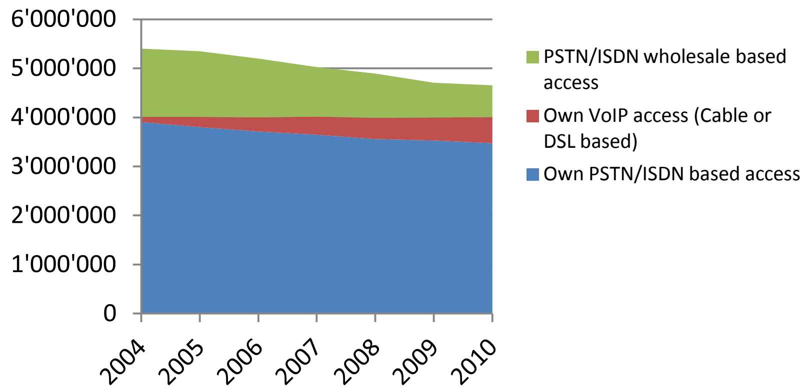


Figure 1 – Number of voice accesses per technology in Switzerland, Source: BAKOM

When looking at fixed voice subscribers per operator (Figure 2) for any technology excluding VoBB, it can be seen that the incumbent continues to have a very high and relatively stable market share of around 65 to 68% from 2007 to 2010, the central period of the period under review. In terms of traffic, it can be seen that Swisscom's market share is consistently about five percentage points lower than in the case of subscribers - at around 60 to 63% (see Figure 3). From this comparison, it seems that some firms (e.g. Sunrise) have customers with longer call duration.

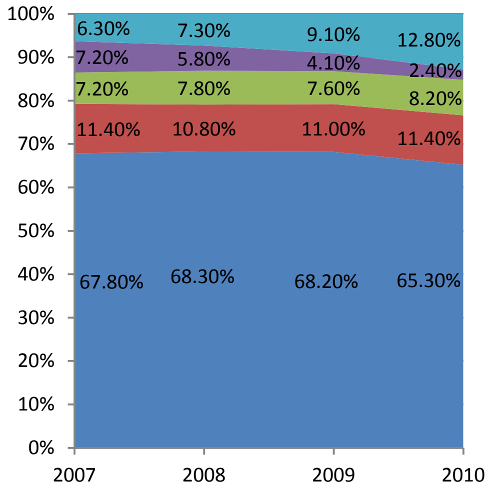


Figure 2 - Fixed voice subscriber market shares in Switzerland including VoIP, excluding VoBB, Source: BAKOM

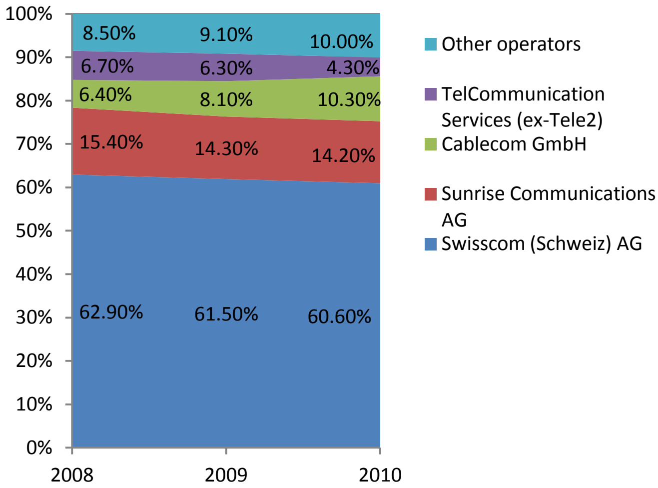


Figure 3 - Fixed voice total traffic markets shares in Switzerland including VoIP, excluding VoBB, Source: BAKOM

In a detailed analysis of the market structure, the traffic data seems to indicate that next to the incumbent having a very high market share, many but significantly weaker competitors exist. The next biggest competitor, Sunrise (copper/DSL based), only holds 14% of subscribers at the end of 2010 (18% of traffic)[27]. Swisscom is therefore about three to four times larger than the next biggest competitor. Moreover, Cablecom (cable based) held 8% of subscribers and 10% of traffic in 2010. The remaining 10% of traffic are distributed among a large number of small competitors of which none exceeded 2% of subscribers or traffic (in total 53 operators were registered Swiss fixed voice operators).

---

[27] This includes former independent competitor Tele2, merged with Sunrise in September 2008.

Of the operators mentioned Sunrise has used only wholesale origination products from the incumbent (carrier selection and pre-selection) until 2008. From 2008, when a regulated unbundling offer was introduced the operator has increasingly migrated to unbundled products and its own VoIP based solutions. Cablecom in turn introduced its cable-based VoIP technology in 2004 (before the operator has not offered telephony services). The other smaller operators may use any of these ways to provide fixed voice services.

3. Theoretical framework

This chapter describes a market model providing the necessary structure for the estimation of (residual) demand elasticities and the degree of competition in the Swiss fixed telephony retail market.

The strong assumptions made in the preceding chapters allow for the design of a simple market model taking into account the volume of national fixed voice traffic (minutes) and corresponding linear retail prices (average revenue per minute). Access charges are assumed to correspond to efficient marginal cost based access prices set by the regulator, implying that both access-seeking operators as well as (assumed efficient) infrastructure-based operators can be assumed to face the same marginal network costs and compete on the same grounds. While mobile termination rates need to be considered as marginal cost drivers, the assumed absence of fixed-mobile substitution and the independence of fixed operators from their potential mobile arms means that fixed voice services and operators are otherwise independent from mobile services and operators (i.e. mobile termination rates are assumed to be exogenous cost drivers). In addition, the relatively low level of regulated fixed termination charges when compared to mobile termination charges mean that when considering a single price for national calls, fixed termination costs and revenues can be ignored. Overall, this setting implies that a market model can be reasonably designed for the Swiss fixed telephony market, which is not significantly more complex than models which would be considered in markets without one- and two-way access aspects. The described simplifications allow for the estimation of a model even with the limited dataset available (2004-2012) described in the next chapter.

The basic framework used in this paper is known as the “dominant firm - competitive fringe” model, and was first proposed by Forchheimer (1908). While this model has relatively strong assumptions, they seem realistic in this specific market. Most importantly, competitors are fragmented and the relative size of the Swiss incumbent fixed telephony operator (60-65%) is clearly above 40%, the threshold for validity of the model indicated by Scherer and Ross (1990). Such a framework is, even under relaxed assumptions, shown in the next chapters to allow the estimation of the residual demand parameters of the incumbent as well as a range for the related conduct parameter measuring the degree of competition. The model is illustrated in Figure 4.

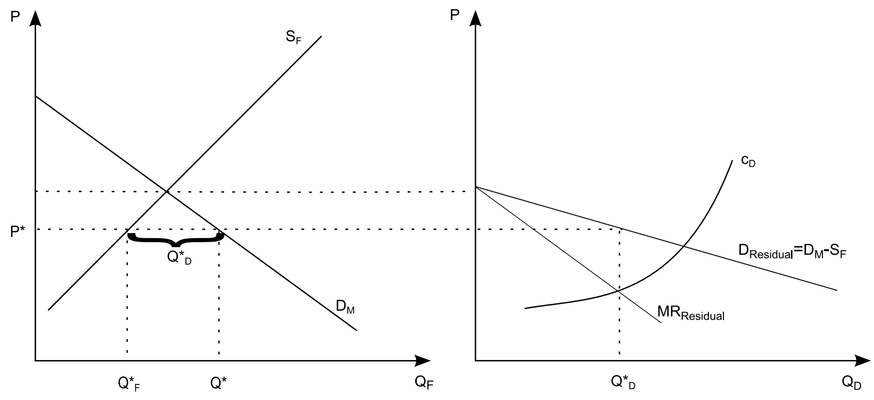


Figure 4 - The dominant firm - competitive fringe model.

Where:

D_M	is the market demand curve (Q is total demand)
S_F	is the fringe supply curve (Q_F is fringe supply)
$D_{Residual}$	is the residual demand curve the dominant firm faces (Q_D is dominant firm demand)
$MR_{Residual}$	is the residual marginal revenue curve the dominant firm faces
C_D	is the dominant firm's marginal costs
P	is the unique market price

In a dominant firm - competitive fringe model, the dominant firm takes into account market demand and fringe supply, and therefore finally the residual demand curve, maximising its profit by equating the resulting perceived marginal revenue to its marginal costs. The residual demand curve faced by Swisscom corresponds to the market demand curve minus the collective supply curve of the supposedly price taking fringe firms. As shown before, it is assumed that there is a largely independent (from mobile) and homogeneous good of national fixed voice traffic (calls to fixed and to mobile networks). A composite unique market price per minute is considered (average revenue per minute).

The dominant firm - competitive fringe model provides the basic structure necessary for the estimation of the residual demand curve and market conduct of the potentially dominant firm. Davis and Garcés (2009) describe a more general setting, with differentiated goods and potentially non-competitive fringe firms going back to Baker and Bresnahan (1988)[28]. Methodologically, the instrumental variable technique with firm specific cost shifters is used to estimate residual demand. This paper proposes a more simplified setting with a homogenous good and a competitive fringe. Practical implementations of the simplified dominant firm – competitive fringe model include Suslow (1986) in the aluminium industry and, more relevant for the present paper, Kahai, Kaserman and Mayo (1996) in the telecommunications industry, where market power of the incumbent telephone operator in the U.S. (AT&T) from 1984 to 1993 is estimated under a dominant firm hypothesis.

In the competitive fringe model it is assumed that the fringe firms are price takers and will – non-strategically - adjust their quantity to the given market price in order to maximize profits. Fringe supply therefore corresponds to the sum of marginal cost curves of fringe firms. The Lerner index measuring market power for the fringe firms is consequently equal to zero (see Lerner (1934)). The dominant firm has a competitive advantage in such a setting; it moves first and takes into account (with perfect foresight) the fringe's reaction and also market demand. In order to estimate the incumbent's residual demand elasticity and its market conduct, the residual demand function has to be estimated. First, in equation (1) market demand is considered.

[28] For homogeneous goods see also Scheffman and Spiller (1987)

$$Q = f(P, X) \tag{1}$$

Here, *Q* is measured by the total number of fixed telephony minutes demanded in the Swiss market (fixed-to-fixed as well as fixed-to-mobile calls), *P* is the average fixed telephony (traffic) price per minute (estimation of average revenue per minute considering all usage based revenues from fixed-to-fixed as well as fixed-to-mobile calls). *X'* is a vector of demand shifting variables. The fringe's marginal cost curve (horizontal sum of single fringe firm's marginal costs[29]) can be modelled in the following form.

$$C_F = f(Q_F, W_F, W_C)$$

Marginal costs are assumed to depend on the fringe's output level and to vary with a series of marginal cost shifters (some specific to the fringe, W_F, and others common to the industry, W_C). The fringe's marginal costs are in the segment under consideration assumed to increase with output due inefficiencies. This, as for example customer care cannot always be easily scaled and as there are capacity constraints of various network elements necessary for the provision of fixed telephony services. The fringe firm's first order condition for quantity choice is $P = C_F$, determining their cumulated supply curve (2).

$$Q_F = f(P, W_F, W_C) \tag{2}$$

The dominant firm, instead, chooses its profit maximizing market price, having consequently monopoly power over residual demand. The first order condition of the dominant firm (in this case it is equivalent for the firm to choose a profit maximizing price or quantity) lead to the following dominant firm supply relation determined by its marginal costs and perceived marginal revenue.

$$P(Q_D) + \frac{P(Q_D)}{\phi_1} = C_D(Q_D, W_D, W_C)$$

Here, ϕ_1, is residual demand elasticity - not to be confused with market demand elasticity - and C_D is the marginal cost function of the dominant firm. W_D represents marginal cost shifters specific to the dominant firm. As in this paper an isoelastic residual demand function[30] will be specified, this implies that the dominant firm supply function can be rearranged to the following expression:

$$P(Q_D) = \frac{C_D(Q_D, W_D, W_C)}{\left[1 + \frac{1}{\phi_1}\right]}$$

Finally, dominant firm's supply can be expressed in log-linear form (lower case variables for natural logs of variables) in **Fehler! Verweisquelle konnte nicht gefunden werden.**

$$p(q_D) = f_D^S(q_D, w_D, w_C) \tag{3}$$

In this case, the constant residual demand price elasticity becomes part of the constant of the function.

As usual, the dominant firm acts as a monopoly on its residual demand, meaning that (in contrast to the competitive fringe firms) it can influence the market price. The more price elastic residual demand is, the lower are prices are and the higher is output.

In this context the Lerner index describes the extent to which the dominant firm can raise prices above its marginal costs:

---

[29] When two identical firms can produce *K* at a given marginal cost, then both firms jointly can produce 2*K*.
[30] i.e. with constant elasticity over the whole range of quantities considered

$$L = \frac{P - C_D}{P} = \frac{1}{\phi_1}$$

Similarly, dominant firm demand, Q_D, can be derived. The dominant firm takes into account market demand (1) as well as fringe supply (2) anticipating perfectly the fringes reactions.

$$Q_D(P) = Q(P, X) - Q_F(P, W_F, W_C)$$

Assuming that residual demand can also be expressed in log-linear form, it can be restated as (4)(see also Scheffman and Spiller (1987), Baker and Bresnahan (1988) and Davis and Garcés (2009)). Such a specification is beneficial in the empirical estimation because coefficients correspond to elasticities and allow for a straightforward discussion regarding percentage changes between variables.

$$q_D = f_D^D(p, x, w_F, w_C) \tag{4}$$

From residual demand (4) the original market demand and fringe supply parameters can usually not be calculated as combined effects are estimated. However, the system of dominant firm demand (4) and supply (3) can now be estimated using standard econometric tools for simultaneous equations. The price can be as usual instrumented using all exogenous variables in the system and be used to estimate the residual demand equation (2SLS). This technique exploits exogenous variation in price allowing for valid regression results. If a technique such as instrumental variables is not used, the assumptions of ordinary least squares regression are violated since the error is then correlated with price.

Before specifying the equations for estimation in detail, it should be noted here that the assumption of a dominant firm with monopoly power is not necessary for the described residual demand approach to hold. In fact, the incumbent can have more competitive conduct than that of a monopolist facing its residual demand and, in the extreme case, even behave as competitively as a price-taker and therefore like the fringe firms. As explained in Bresnahan (1982), different strategic conduct by the incumbent can be nested in its profit maximizing condition by simply adding a *conduct parameter,* λ, in the marginal revenue function (see also Davis and Garcés (2009). The approach to take the strategic conduct of firms as a parameter to be estimated and not as a model assumption is the core of the so-called “new empirical industrial organization” described in Bresnahan (1989). Including a conduct parameter λ for the potentially dominant firm, the above optimality conditions can be restated as

$$P(Q_D) + \lambda \frac{P(Q_D)}{\phi_1} = C_D(Q_D)$$

$$P(Q_D) = \frac{C_D(Q_D, W_D, W_C)}{\left[1 + \frac{\lambda}{\phi_1}\right]}$$

Here λ takes values between zero and one. $\lambda = 1$ corresponds to the working hypothesis of this paper of a dominant firm. In this case the incumbent behaves as a monopolist facing its residual demand. This setting is relaxed with $\lambda < 1$. In the extreme case, $\lambda = 0$. Then, the incumbent would be a price-taker without market power, as all its competitors. Any value between zero and one would correspond to the level of market power exercised (adjusted by elasticity). In particular, the Lerner index can be adapted to the following equation:

$$L = \frac{P - C_D}{P} = \frac{\lambda}{\phi_1}$$

As λ may be considered as a given parameter for the incumbent over the time horizon under review, the earlier derived incumbent’s supply and residual demand functions remain valid if the incumbent does not behave fully as a dominant firm. The same is true for the instrumentation.

It should be noted that Bresnahan (1982) shows how to estimate λ in linear demand and marginal cost settings by using demand (and cost) rotating instead of shifting variables. In simpler settings, as in the context of this paper, λ can often not be directly estimated. For some particular values of the residual demand price elasticities, however, inferences on λ are possible. In particular, when residual demand is inelastic a maximal λ can be estimated for which the incumbents' action are compatible with profit maximizing behavior (i.e. perceived marginal revenue being positive). Concretely, it is shown in the next chapters that in light of the estimation results obtained (inelastic residual demand at -0.12), λ must be below one (and in particular below 0.12) to be compatible with profit maximization. The incumbent does therefore not act as a (purely) "dominant" firm. Nevertheless, the assumption of the incumbent acting as a dominant firm is for now maintained as a working hypothesis[31].

First model equation (instrumented price):

As has been shown, considering all exogenous variables in the system, the market price can be instrumented. The system includes market demand, fringe supply and dominant firm supply. This means that demand shifters, common marginal cost shifters, fringe specific marginal cost shifters and, most importantly, dominant firm specific cost shifters need to be used as instruments to fit a variable $\hat{p}$ in (5).

$$\hat{p} = \theta_0 + w_D'\theta_D + w_F'\theta_F + w_C'\theta_C + x'\theta_B + \varepsilon \qquad (5)$$

Second model equation (dominant firm residual demand)

The instrumented price can then be used to estimate the residual demand function of the (potentially) dominant firm (6).

$$q_D = \phi_0 + \phi_1\hat{p} + w_F'\phi_F + w_C'\phi_C + x'\phi_B + \varepsilon \qquad (6)$$

The coefficients of this function are estimated in the next chapter and allow inferences to be made on residual demand elasticity and market power. As has been shown before, the residual demand (price) elasticity is the result of a combination of market demand and fringe supply effects (see to (4)). Regarding the original individual coefficients of market demand and fringe supply, these can in general not be derived from the residual demand estimates in this context.

Some papers estimate fringe supply and market demand exclusively and make inferences on residual demand elasticity without direct estimation of the residual demand function. Kahai, Kaserman and Mayo (1996) use this approach as they argue that there are no sufficiently strong cost shifters for the dominant firm in order to estimate residual demand. While this may formally be a correct approach, the model applied by these authors largely abstracts from the dynamics of the most important factor of competition in this model; the potentially dominant firm. In the present paper, the marginal cost shifters of the incumbent are assumed to be important and are taken into account as instruments (e.g. staff numbers of Swisscom). Not taking these variables into account could cause a bias in the model independently of how it is specified. The model proposed here therefore allows for convenient direct estimation of the residual demand function and corresponding elasticities.

4. Input data

This chapter describes the dataset used to estimate the market model.

[31] It should be noted that the model remains also broadly valid when the assumption of a competitive fringe is relaxed (see Bresnahan and Baker (1988)).

In the dataset considered in this paper, different quarterly time series to model competition in the fixed voice market in Switzerland from 2004 to 2012 are used. The strategic variables include the dominant firm's traffic output, i.e. Swisscom's fixed national outgoing calling minutes (to fixed and to mobile) and an estimate of market prices (average revenue per outgoing minute (subscription revenues excluded). The total market includes all fixed telephony technologies (e.g. PSTN as well as proprietary VoIP), but excludes web-based Voice over IP, as there is no reliable data on usage and prices and substitutability is assumed to be yet limited.

Traffic demand shifters include real GDP per capita, the network size (number of active fixed telephone lines in the industry, assumed to increases the number of potential calls) and time dummies. As it is assumed here that access is largely independent from traffic, the total network size is assumed to be exogenous, as in Kahai, Kaserman and Mayo (1996). Moreover, a series of common industry marginal cost shifters are considered, which include most importantly the average mobile termination rates, the actual regulated origination rates (as a proxy for a part of the network cost), interest rates and exchange rates (as a proxy of capacity investment costs). In addition, fringe cost shifters include the extent of usage of ULL and ADSL wholesale broadband products by alternative operators (essentially for a question of economies of scope reducing for example per unit customer care cost). Finally, cost shifters specific to Swisscom are considered, which include the number of staff (to control for possible remaining inefficiencies from times when the firm was fully state controlled). In addition, the number of ADSL lines sold by the incumbent may lead to economies of scope in a similar way as for the alternative operators. The effects described here are analysed in more detail in a later chapter interpreting the estimation results.

For the variables discussed, quarterly observations are available from Q4 2004 to Q2 2012. The variables in Table 1 are used to model the Swiss fixed voice market in the next chapter.

Variable	Definition	Name in Stata dataset	Unit	Source	N	Mean (abs)	Std. dev. (abs)
p	Price; average revenue per outgoing minute (subscription revenues excluded), deflated by CPI (100=2006), traditional and proprietary VoIP	*tradvoipallar pmallr*	CHF	Analysys Mason, BFS	55	0.098	0.017
q_D	Swisscom's fixed national outgoing calling minutes (to fixed and to mobile)	*tradvoipallmi nscm*	Minutes	Swisscom	55	2.45E+09	5.24E+08
Demand shifters							
x_1	Income; real GDP per capita (100=2006)	*yrealcapita*		Seco / BFS	51	0.013	0.0006
x_2	Number of active telephony lines (PSTN)	*pstn*	000s	Swisscom	56	2'693	1'071
x_3, x_3, x_4	Quarterly dummies for quarters 2, 3 and 4	*d2, d3, d4*					
Fringe supply shifters							
w_{F1}	Number of wholesale ADSL lines sold by Swisscom	*adslwhole*	000s	Swisscom	47	262.996	147.691
w_{F2}	Number of ULL access lines sold by Swisscom	*ullreal*	000s	Swisscom	59	53.288	103.362
Common supply shifters							
w_{C1}	Average regional fixed voice origination prices per minute deflated by CPI (100=2006)	*scmregorr comcregorr*	CHF (100= 1/2000)	Swisscom /ComCom	55	0.012	0.004
w_{C2}	Weighted average of mobile termination rates deflated by CPI (100=2006)	*wavgmtrr*	CHF	Operators Analysys Mason, BFS	34	0.21	0.089
w_{C3}	Interest rates on 30y bonds of the Swiss Confederation	*interest30y*	%	Swiss national bank	56	3.19	1.026
w_{C4}	Exchange rate EUR/CHF	*fxratechfeur*	EUR/CHF		56	0.676	0.067
Dominant firm supply shifters							
w_{D1}	Number of staff working for Swisscom (Group)	*SCMstaff*	units	Swisscom	35	18418	1880.025
w_{D2}	Number of active retail ADSL lines (Swisscom)	*adslretail*	000s	Swisscom	47	885	602

Table 1 – Input data from the Swiss fixed telephony market (2004-2012)

It should be noted that for the estimation natural logs of all variables are taken (the Stata variables names in the technical annex in this case have an "ln" prefix[32].

[32] Having only about 30 observations, it is early to estimate this market model. A later estimation could take advantage of more data points. Standard errors seem to be high, which is often the case in small samples.

5. Econometric analysis and estimation

This chapter reports the estimation results. As is typical in a demand and supply setting, instrumental variables can be used for estimation. The basic econometric model can be analysed in two stages as described in the analytical framework. In the technical annex (Section 9.1.) a "baseline" or reference model in line with the analytical framework is described and estimated. It is shown that important adjustments to the basic model are necessary in order to resolve the detected econometric problems. Most importantly, it is found that fixed voice traffic and prices are decreasing over time, implying non-stationarity both for the incumbent's and the fringe's strategic variables. In addition, errors are serially correlated and strategic variables seem to be autocorrelated (most importantly with their first lag). The technical annex analyses these problems and develops the solution adopted in this section in the form of an Auto-Regressive Distributed Lags (ARDL[33]) model. It specifically implies that the baseline equations are added one period lagged dependent and independent variables (ARDL (1,1)). This adaption is necessary to exclude spurious regression results as a consequence of non-stationarity and insufficient cointegration of variables. As the model becomes dynamic, the interpretation of specific coefficients becomes more complex. While coefficients continue to correspond to short run (same period) direct effects on the dependent variable ("impact multipliers"), the long term effects of a (permanent) change in an explanatory variable need to be calculated ("long run multipliers", see Equation (9) in Section 9.7. of the technical annex).

The technical annex overall concludes that a first stage ARDL (1,1) regression is estimating an instrumented variable with instruments (all exogenous variables in the system) that are non-stationary, but having a sufficient degree of cointegration. Similarly, the second stage ARDL (1,1) regression is estimated with variables that are non-stationary, but having a sufficient degree of cointegration (including the instrumented variable). In addition, in both regressions with these specifications, there is no serial correlation of errors anymore. Overall, both the first and second stage ARDL regressions are valid. Hence, in the next section, the ARDL (1,1) regressions are used to correct the baseline model (see technical annex) for the econometric problems identified, especially spurious regression.

Estimation

All necessary tests to exclude spurious regression results have been performed in the technical annex. The following ARDL (1,1) second stage regression (following a similar first stage regression) should therefore represent the optimal 2SLS model in the context of this paper and provide valid estimation results for the incumbent's residual demand function. For convenience, Equation (12) from the technical annex specifying residual demand of the incumbent may be restated here:

$$q_{D,t} = \hat{\phi}_0 + \hat{\phi}_{1,t}\hat{p}_t + \hat{\phi}_{1,t-1}p_{t-1} + \hat{\phi}_{2,t-1}q_{t-1} + w'_{F,t}\hat{\phi}_{F,t} + w'_{F,t-1}\hat{\phi}_{F,t-1} + w'_{C,t}\hat{\phi}_{C,t} + w'_{C,t-1}\hat{\phi}_{C,t-1} + x_t{}'\hat{\phi}_{B,t} + x_{t-1}{}'\hat{\phi}_{B,t-1} + \epsilon_t \tag{12}$$

Estimation results are reported in Table 2, where the coefficients (impact multipliers) of the ARDL regression are represented next to the baseline model coefficient estimates. The latter are, however, as shown in the annex, likely to be spurious. The dynamic multipliers (impact on following period) are represented one row below (L1). Finally, another row below long run multipliers, $\hat{\phi}^{LR}$, are calculated, which represent the effect of a change in an independent variable over the whole time horizon on the dependent variable. The last column compares the coefficient estimates of the ARDL(1,1) estimation with the baseline estimation. It should be noted, however, that from the tests conducted the ARDL model is clearly the correct model for estimation and the baseline model results are only reported for convenience.

[33] Or also ADL

Parameters	Variable	ARDL (1,1) estimates: Coefficient estimate	Robust Std. err.	P>|t|	Baseline estimates: Coefficient estimate	Robust Std. err.	P>|t|	Comparison: Difference ADLR/Baseline (to impact and LR multplier) in %
	q_D							
$\hat{\phi}_{2,t-1}$	L1	-0.181	0.100	0.069				
$\hat{\phi}_{1,t}$	p	-0.044	0.044	0.321	-0.663	0.364	0.068	-93%
$\hat{\phi}_{1,t-1}$	L1	-0.102	0.049	0.038				
$\hat{\phi}_{1}^{LR}$		-0.124						-81%
$\hat{\phi}_{B1,t}$	GDP per capita	1.364	0.302	0.000	-0.333	0.726	0.646	-509%
$\hat{\phi}_{B1,t-1}$	L1	0.387	0.236	0.102				
$\hat{\phi}_{B1}^{LR}$		1.482						-545%
$\hat{\phi}_{B2,t}$	PSTN lines	0.876	0.159	0.000	1.979	0.366	0.000	-56%
$\hat{\phi}_{B2,t-1}$	L1	1.912	0.337	0.000				
$\hat{\phi}_{B2}^{LR}$		2.360						19%
$\hat{\phi}_{B3,t}$	d2	-0.030	0.005	0.000	-0.032	0.016	0.042	-5%
$\hat{\phi}_{B4,t}$	d3	-0.098	0.006	0.000	-0.065	0.024	0.007	51%
$\hat{\phi}_{B5,t}$	d4	-0.025	0.010	0.008	-0.021	0.013	0.112	23%
$\hat{\phi}_{F1,t}$	ADSL wholesale lines	-0.233	0.058	0.000	-0.220	0.043	0.000	6%
$\hat{\phi}_{F1,t-1}$	L1	0.046	0.054	0.391				
$\hat{\phi}_{F1}^{LR}$		-0.158						-28%
$\hat{\phi}_{F2,t}$	ULL lines	-0.003	0.001	0.011	-0.001	0.001	0.548	315%
$\hat{\phi}_{F2,t-1}$	L1	0.000	0.001	0.700				
$\hat{\phi}_{F2}^{LR}$		-0.002						196%
$\hat{\phi}_{C1,t}$	Origination prices	-0.097	0.026	0.000	0.150	0.133	0.259	-165%
$\hat{\phi}_{C1,t-1}$	L1	0.037	0.026	0.156				
$\hat{\phi}_{C1}^{LR}$		-0.051						-134%
$\hat{\phi}_{C2,t}$	MTR	0.010	0.021	0.621	0.116	0.071	0.104	-91%
$\hat{\phi}_{C2,t-1}$	L1	0.157	0.018	0.000				
$\hat{\phi}_{C2}^{LR}$		0.141						22%
$\hat{\phi}_{C3,t}$	Interest rate	-0.112	0.034	0.001	-0.050	0.052	0.338	123%
$\hat{\phi}_{C3,t-1}$	L1	0.070	0.025	0.005				
$\hat{\phi}_{C3}^{LR}$		-0.035						-30%
$\hat{\phi}_{C4,t}$	Exchange rate	0.197	0.141	0.163	-0.220	0.155	0.155	-189%
$\hat{\phi}_{C4,t-1}$	L1	0.362	0.149	0.015				
$\hat{\phi}_{C4}^{LR}$		0.473						-315%
$\hat{\phi}_{0}$	_cons	11.038	1.928	0.000	4.390	4.102	0.285	151%
$\hat{\phi}_{0}^{LR}$		9.343						113%
		R^2	0.998		R^2	0.966		
		Prob> χ^2	0.000		Prob>F	0.000		
		N	27		N	28		

Table 2 – Estimation results, Baseline and ARDL(1,1)[34]

As has been shown, the results of the ARDL(1,1) regression should not be the result of a spurious regression and can be interpreted as usual in the next section.

---

[34] Stata command: ivregress 2sls lntradvoipallminscm l.lntradvoipallarpmallr l.lntradvoipallminscm d2 d3 d4 l(0/1).lnyrealcapita l(0/1).lnpstn l(0/1).lnscmregorr l(0/1).lnadslwhole l(0/1).lnullreal l(0/1).lnwavgmtrr l(0/1).lninterest30y l(0/1).lnfxratechfeur (lntradvoipallarpmallr=l(0/1).lnSCMstaff l(0/1).lnadslretail) , first robust

6. Interpretation of the results

In this chapter the main estimation results are discussed. After performing a detailed analysis of the stability of the baseline results and introducing the necessary corrections to account for cointegration and serial correlation of errors (see technical annex), the resulting autoregressive distributed lags model regression results (Table 2) can be interpreted as usual 2SLS results. This section shows how, as for any model which includes intertemporal effects, the effect of a change in one variable on another variable has to be divided into an immediate same period effect (impact multiplier) and a long term effect (long run multiplier). It is possible to use the estimated coefficients to calculate functions illustrating the intertemporal adjustment behaviour of the dependent variable after exogenous shocks over time.

For convenience (9) from the technical annex can be restated here:

$$y_t = c + \alpha_1 y_{t-1} + \beta_0 z_t + \beta_1 z_{t-1} + \epsilon_t$$

From the explanations in the technical annex (see (9) and (10)) it follows that the time t+k multiplier (i.e. the coefficient explaining the impact of a marginal change of an independent variable in period t on the dependent variable from t until t+k) is described by the following expressions:

$$\mu_{t+k} = \frac{\partial y_t}{\partial z_t} + \frac{\partial y_{t+1}}{\partial z_t} + \frac{\partial y_{t+2}}{\partial z_t} + ..$$

$$= \beta_0 + (\alpha_1\beta_0 + \beta_1) + \alpha_1(\alpha_1\beta_0 + \beta_1) + \alpha_1^2(\alpha_1\beta_0 + \beta_1) + ...$$

If $k = 0$ $\quad \mu_{t+k} = \beta_0$

If $k \neq 0$ $\quad \mu_{t+k} = \beta_0 + \sum_{k=1}^{t} \alpha_1^{k-1}(\alpha_1\beta_0 + \beta_1)$

If $k \to \infty$ $\quad \mu_{t+\infty} = \sum_{k=0}^{\infty} \alpha_1^{k}(\beta_0 + \beta_1)$ (7)

The multiplier with k=0 is the called the impact multiplier.

As long as α_1<1, which is assumed here[35], the long term multiplier defining the effect on the steady state of the dependent variable from a change in an independent variable in period t corresponds to (8) (Johnston and Di Nardo (1997)).

$$\frac{dy^*}{dz_t} = \frac{\beta_0 + \beta_1}{1 - \alpha_1} \qquad (8)$$

These functions, representing the short, medium and long run multipliers, may in the present log-linear model also be interpreted as “cumulative” elasticities for a period up to k periods after the assumed shock in the independent variable of interest. Calculating these effects, Figure 5 represents the percentages of full adjustment after k quarters after the shock, where the full adjustment corresponds to the long run multiplier. This contrasts with the baseline regression, where it is assumed by definition that 100% of the adjustment is taking place in the same period as the shock. It can be seen that residual demand in the estimated model nearly fully adjusts to shocks in the market in all cases after only four quarters. A large part of the full adjustment for most shocks takes place in the same quarter as the shock as well as in the following quarter. It can also be noted here that for several variables there seems to be some form of overshooting effect in the quarter of the shock .

[35] otherwise effects on the dependent variable over time become ever stronger

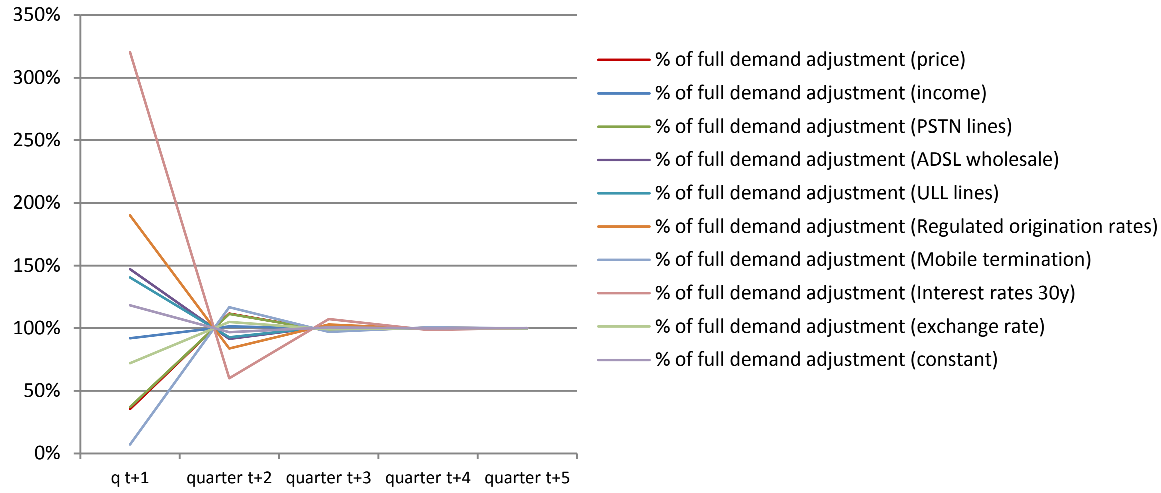


Figure 5 – Percentage of long term adjustments in the first quarters after the shock

Considering the above, the interpretation of results can be limited to effects in the first four quarters after the shock and most importantly the period of the shock and the following period. The most important estimated cumulative residual demand elasticities for the incumbent are reported in Figure 6.

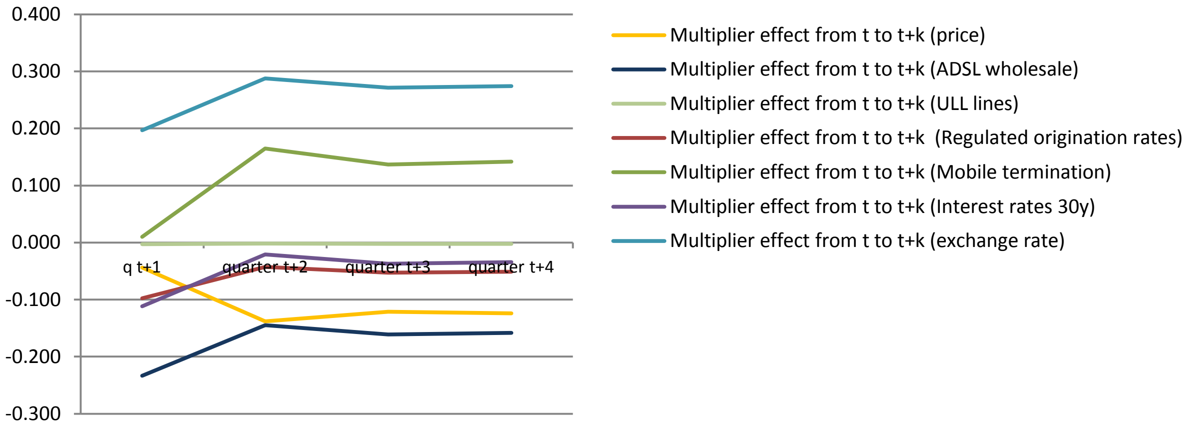


Figure 6 – Main estimated cumulative residual demand elasticites up to period t+k

In the following sections, the single coefficient estimation results and corresponding cumulative elasticities are discussed.

6.1. Price elasticity of residual demand

For convenience, single coefficient results of Table 2 are reported again in the corresponding subsections before interpreting the coefficients.

Variable	ARDL (1,1) estimates			Baseline estimates			Comparison
	Coefficient estimate	Robust Std. err.	P>| t |	Coefficient estimate	Robust Std. err.	P>| t |	%
p	-0.044	0.044	0.321	-0.663	0.364	0.068	-93%
L1	-0.102	0.049	0.038				
LR	-0.124						-81%

Table 2.1. – Estimation results, coefficient on price

It can be seen that the elasticity of residual demand of Swisscom in response to a price change is very low during the quarter in which prices are adjusted (-0.044). The direct effect on the period following the shock is higher, at -0.102. Finally, compared to the same period effect, the long term cumulative elasticity, taking into account all of the price change on residual demand over subsequent periods, triples to -0.124.

It may be seen as unusual in as fast paced an industry as telecommunications that a large part of the demand adjustment takes several months (or even quarters) to materialize. The contractual terms of Swisscom (and others operators) in the market during the period under review may, however, explain this phenomenon. A reduction in voice minutes demanded may come from users calling less or from users fully giving up fixed telephony services. In the period under review, Swisscom required that its private customers provide a 60 days' notice period to cancel the service. During the first 60 days after the adjustment, volume may therefore only be affected by customers calling less. Afterwards, it will also be affected by customers fully giving up their fixed access lines (or switching). This means that for two-thirds of a period after the shock a large part of the adjustment is blocked and possibly compensated only subsequently[36].

Nevertheless, even taking into account the long term effects of a price change for fixed voice traffic, residual demand of Swisscom seems to be highly inelastic (-0.124). This indicates that Swisscom, acting as a dominant firm, could increase prices without fearing immediate and even medium and long run material residual demand adjustments. In particular, after a 10% price increase, demand for Swisscom fixnet traffic would decline by only 0.4% in the same quarter, 1% in the following quarter (direct impact only) and 1.2% cumulated in the long term.

As has been shown earlier for any type of demand function, the marginal revenue of a dominant firm is given by

$$MR = P(Q)\left[1 + \frac{1}{\phi_1}\right]$$

Where ϕ_1 is estimated at -0.124. When this price elasticity is lower (in absolute terms) than 1, marginal revenue is by definition negative. As on the other hand marginal costs are always positive, a profit maximizing dominant firm would in such a situation increase prices and reduce its output up to a point where residual demand becomes elastic and where MC=MR. This is not directly possible in the model proposed, as the log-linear form of residual demand assumes constant price elasticity. This is, however, a simplifying assumption (see Davis and Garcés (2009)). If the incumbent would actually increase prices, this would in reality possibly lead an increased price elasticity and the model would also find a higher (constant) level of price elasticity. In any case, the model and its estimates indicate that at this outcome Swisscom cannot be a profit maximizing dominant firm. Graphically, the estimated functions can be illustrated in Figure 7.

[36] It should be noted that, in the particular case of standard rates, upwards price adjustments are often notified in advance, implying that customer reactions in this case may be more immediate.

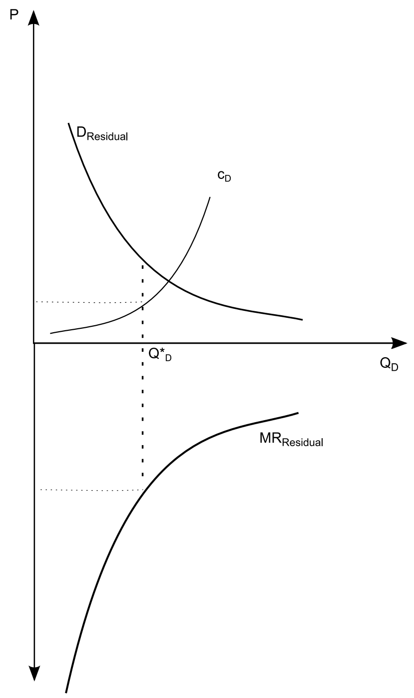


Figure 7 – Estimated residual demand curve and marginal revenue (dominant firm)

As also explained earlier, the market model is also compatible with the incumbent acting strategically differently, i.e. as a price taker (no market power) or between a price taker and a monopolist facing residual demand. The marginal revenue becomes:

$$MR = P(Q)\left[1 + \frac{\lambda}{\phi_1}\right]$$

For Swisscom to be able to behave in a profit maximizing manner, marginal revenue would need to be positive. With ϕ_1=-.12, this is the case only when $\lambda < 0.12$. In this case, the marginal revenue curve would as usual be positive and a profit maximizing firm could produce at the intersection point with marginal cost (see Figure 8).

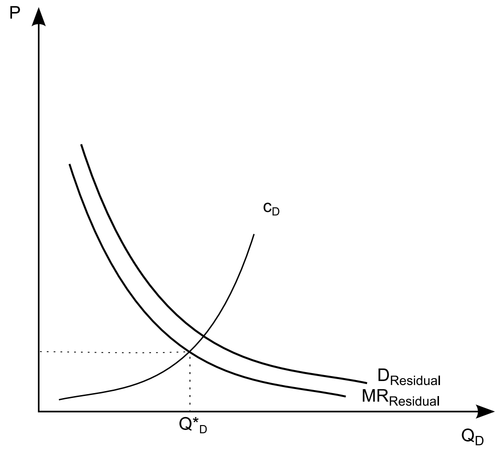


Figure 8 - Estimated residual demand curve and marginal revenue (not fully dominant firm; $\lambda < 0.12$)

This implies that only when Swisscom acts not as a dominant firm but nearly as competitively as the fringe firms, its actions are compatible with the standard assumption of profit maximization. It can therefore be concluded that Swisscom's conduct is largely competitive. To cite one example, Qu (2007) finds even lower price elasticity of residual demand, not significantly different to zero for strategic firms in the wholesale electricity market in the U.S. He similarly concludes that this implies that the incumbent's behaviour is consistent with fully competitive pricing. It should be noted that this is not the case in this paper. While the immediate, same period effect (-0.044) is also not significantly different than zero, subsequent effects are. The estimates in this paper imply only that Swisscom must act rather competitively. It can, however, not be affirmed whether or not the firm acts fully competitively.

It should also be noted, that for a small part of the considered retail telephony price basket, binding price regulation via universal service exists (fixed-to-fixed calls for standalone offers). For these prices, the fact that Swisscom is not increasing prices as a dominant firm while facing such inelastic demand may be not related to competition but to regulation. It can therefore not be said a priori for these prices that lifting regulation would not lead to increased prices, as currently regulated prices distort this analysis. These regulated prices are, however, only of limited validity as a large number of Swisscom price plans and subscriptions over the period under review foresee non standard per minute rates. For the composite good under analysis (the incumbent's fixed retail voice traffic overall) the conclusion should therefore hold that the incumbent behaves largely competitively and that any price regulation of this market is unnecessary and potentially distorting competition.

6.2. Demand shifters

Variable	ARDL (1,1) estimates			Baseline estimates			Comparison
	Coefficient estimate	Robust Std. err.	P>| t |	Coefficient estimate	Robust Std. err.	P>| t |	%
Income	1.364	0.302	0.000	-0.333	0.726	0.646	-509%
L1	0.387	0.236	0.102				
LR	1.482						-545%
PSTN lines	0.876	0.159	0.000	1.979	0.366	0.000	-56%
L1	1.912	0.337	0.000				
LR	2.360						19%

Table 2.2. – Estimation results, coefficients on demand shifters (income, PSTN lines)

Income

According to the estimates an increase of per capita income of 10% is expected to correspond to an increase in fixed voice traffic demand for the incumbent of 14.8%. This effect may be slightly stronger than expected, with an income elasticity higher than one, meaning that fixed voice traffic in this context is suggested to be superior good.

While studies usually find that fixed voice traffic is a normal good (e.g. Ahn, Lee and Kim (2002) for Korean fixed telephony), there are also studies supporting the view of a superior good, as for example, Gyimah-Brempong and Karikari (2007) for African countries. Agüero, De Silva and Kang (2011) similarly review Engel curves for (essentially fixed) telecommunications services and find that early studies and studies on developing countries mostly found fixed telephony services to be a luxury good, while in developed countries these would usually be rather normal goods. McCloughan and Lyons (2006) review income effects on mobile telephony services which they find to be usually a normal or superior good. They also state that the income elasticity of demand may depend on the proportion of high income customers served. Vogelsang (2010) overall shares these views stating that mobile telephony may be a luxury good initially but may become a normal good in a more mature phase.

It may seem reasonable that demand for (more expensive) mobile calls increases more strongly with an increase in income than demand for fixed calls (the cheaper and less convenient alternative). Nevertheless, it also seems reasonable that both goods in the period under examination were perceived as superior goods in Switzerland, as unregulated mobile termination rates and telephony prices from any network towards mobile were as shown extraordinarily high.

PSTN lines

The model suggests also that there are positive network effects in the sense that the larger the fixed network (PSTN lines[37]) the larger the amount of persons that can be reached from fixed accesses via particularly cheap fixed-fixed calls suitable for longer calls. As also in Kahai, Kaserman and Mayo (1996) the total network size is assumed to be exogenous and therefore largely independent from traffic (standalone fixed access prices of all operators have for instance remained constant in the period under analysis and do not vary across operators). It should be noted here that the PSTN lines variable is also representing the effects of growth in population. This variable has actually been dropped due to strong colinearity with PSTN lines.

6.3. Fringe cost shifters

Variable	ARDL (1,1) estimates			Baseline estimates			Comparison
	Coefficient estimate	Robust Std. err.	P>| t |	Coefficient estimate	Robust Std. err.	P>| t |	%
ULL lines	-0.003	0.001	0.011	-0.001	0.001	0.548	315%
L1	0.000	0.001	0.700				
LR	-0.002						196%
ADSL wholesale lines	-0.233	0.058	0.000	-0.220	0.043	0.000	6%
L1	0.046	0.054	0.391				
LR	-0.158						-28%

Table 2.3. – Estimation results, coefficients on fringe cost shifters (ULL lines, ADSL wholesale lines)

The level of ADSL wholesale lines sold by the incumbent (to the fringe firms) and the level of usage of unbundling (ULL) are proxies for the efficiency with which fringe telephony operators can operate their central infrastructure and services (while regional origination costs are assumed to be the same for all operators). Efficiencies are mainly expected due to economies of scope by providing both telephony and Internet services.

The Internet market is assumed to be largely independent from the telephony market here, which is why the ADSL wholesale lines are taken as exogenous. The level of unbundling needs not necessarily be correlated with broadband penetration (see Nardotto, Valletti and Verboven (2012)). However, unbundling can in this market model be expected to mainly indicate the technological state of the fringe firms' backbone network, as for a large part of the period under review ULL was not available and for a subsequent relatively long period the ULL infrastructure was only successively rolled out starting with the most profitable areas. It can therefore be expected that the market environment (such as wholesale and retail prices) could have affected the level of demand for ULL only for a small part of the period under review.

The results show, however, that the efficiency effect by ULL in the backbone network seems to be limited. The number of (wholesale) broadband lines provided by the fringe firms seem, however, to have a more substantial cost reducing effect expanding as expected fringe supply and limiting residual demand.

---

[37] VoIP lines were largely irrelevant for large part of the period under review.

6.4. Common cost shifters

Variable	ARDL (1,1) estimates			Baseline estimates			Comparison
	Coefficient estimate	Robust Std. err.	P>| t |	Coefficient estimate	Robust Std. err.	P>| t |	%
Mobile termination rates	0.010	0.021	0.621	0.116	0.071	0.104	-91%
L1	0.157	0.018	0.000				
LR	0.141						22%
Origination prices	-0.097	0.026	0.000	0.150	0.133	0.259	-165%
L1	0.037	0.026	0.156				
LR	-0.051						-134%
Interest rate	-0.112	0.034	0.001	-0.050	0.052	0.338	123%
L1	0.070	0.025	0.005				
LR	-0.035						-30%
Exchange rate	0.197	0.141	0.163	-0.220	0.155	0.155	-189%
L1	0.362	0.149	0.015				
LR	0.473						-315%

Table 2.4. – Estimation results, coefficients on common cost shifters (MTR, origination, interest and exchange rate)

The interpretation of common cost shifters has to be done carefully. A common positive cost shock affects both the fringe firms as well as the dominant firm. The above coefficients measure, however, only the impact via fringe supply on residual demand. The impact on incumbent supply is not measured and not the focus of this paper.

Mobile termination rates

The voice traffic considered includes fixed to mobile calls. In the Swiss setting, with unregulated (and relatively high) mobile termination rates in the period under review, these probably are the most important cost drivers for national traffic (fixed termination costs and revenues continue to be abstracted from here). As explained earlier, these rates are assumed to be exogenous. In such a case, it can be expected that any drop in the average mobile termination rate would have a positive effect on fringe output reducing dominant firm demand. It is found that an overall decrease of 10% in mobile termination rates would reduce incumbent demand by 1.4% in the long term. Again, incumbent equilibrium output may differ as the effect of these common costs on incumbent supply would also need to be considered.

Exchange rates

Telecommunications equipment is typically supplied by firms outside Switzerland. When the strength of the local currency increases it can be expected that telecoms equipment may be sourced at lower prices. The Euro area is by far the largest import area for the Swiss economy, the EUR/CHF is therefore considered. An increase in the exchange rate EUR/CHF (around 0.8 in 2012) would imply that more Euros are obtained per Swiss Franc. In this case the purchasing power of Swiss companies should at international level increase and sourcing prices decrease. This effect should be common for all firms in the market. The effect on fringe firms would be an expansion of supply and therefore a reduction in dominant firm demand. This effect is not seen in the results, however. In fact, the same period effect is insignificant, while the one quarter lagged effect is slightly positive. It is therefore likely that such equipment is not a relevant marginal cost driver in the fixed voice market.

Interest rates

Another common cost driver is capital costs proxied by the interest rate (Swiss 30-year Government Bond). An increase in capital costs is expected to lead to a reduction of fringe supply and an increase in the incumbents' residual demand. Here the lagged effect is positive as expected but the same period and the long term effects are negative. Again, in this market marginal costs may be only partially affected by significant capital investment.

Finally, it may be noted that the state-owned incumbent may be less dependent on the capital market. This is, however, a supply side consideration which is not directly considered here.

Origination rates

Incurred marginal costs for network origination are as explained earlier assumed to be the same for all operators (infrastructure-based operators as well as access-seekers). Regulated origination rates are used here as a proxy for (regional) network origination cost for all operators.

An increase of origination rates for providers for fixed voice calls should increase all operators' origination cost and reduce fringe output increasing Swisscom's residual demand. While this is again the case for the one period lagged effect, this is not the case for the same period as well as the long term effect. This ambiguous effect may mean that origination rates may be of limited importance on the retail market. This may be reasonable given their comparatively low level when compared to mobile termination rates (per minute).

7. Conclusions

The simple simultaneous equations framework proposed in this paper, based on a generalised "dominant firm – competitive fringe" model has allowed to estimate the Swiss incumbent's residual demand function for fixed telephony traffic for the period from 2004 to 2012. Unlike earlier papers, this paper directly estimates residual demand using dominant firm specific cost shifters and ensures a sufficient level of cointegration to avoid spurious regression results. Evidence of a competitive fixed telephony market in Switzerland during the period under review is found, calling into question the need for continued regulation.

While conduct cannot be directly estimated using the framework described, the concrete estimates show that demand is inelastic (long run price elasticity of -0.12). Such a level of elasticity is, however, only compatible with a profit maximising incumbent in the case of largely competitive conduct (conduct parameter below 0.12 and therefore close to zero). It is therefore found that Swisscom acted rather competitively in the fixed telephony retail market in the period under review. If the problem of an uncompetitive retail market ever existed, it seems that the entry of alternative operators using cable infrastructure and the introduction of regulated wholesale access (carrier pre-selection as well as local loop unbundling) have successfully removed it. This implies that the (partial) retail price caps in place in Switzerland can no longer be justified on the basis of a lack of competition and should be removed as has been done recently in a number of other European states. Similar conclusions can be drawn for possible universal service objectives. As regulated wholesale products (call origination) are available on national scale at a uniform price, retail competition should be ensured at national scale as well and not only locally.

The model in this paper is based on a series of strong assumptions. Future empirical work should try to relax them. Most importantly, the model assumes that all operators face the same marginal network cost. In reality, this may not be the case. The incumbent may face lower internal marginal network costs in some form. In this case, vertical integration of the incumbent becomes relevant. As has been discussed, Inderst and Peitz (2012) show then that in case of price dependent demand[38], the incumbent could also charge lower uniform retail prices in equilibrium than its competitors and that it then has a higher market share (partial foreclosure). Such pricing policy corresponds to a margin squeeze by the incumbent (see Vickers (2005)). If such a margin squeeze has occured, competitive conduct by the incumbent in the retail market could be overestimated. Different firm-specific marginal network costs and prices (differentiated goods) should therefore be considered. More generally, the market for inputs should be modelled in more detail in future work. This could be of particular interest, as the current regulatory debate on European level is starting to focus on these markets. In particular, an Ecorys study for the European Commission (2013) concludes that regulated unbundling and wholesale broadband access products sufficiently constrain the incumbent in its call origination pricing in the future, suggesting it may be possible to lift also the regulation of call origination[39]. It is possible that this paper's focus on this regulated access product was inappropriate, even retrospectively, and may have caused the unexpected sign of the related coefficient as discussed in the last chapter. Whether the European Commission shares the view of Ecorys remains to be seen. An updated recommendation on relevant markets is due in the second quarter 2014.

Lastly, fixed-mobile substitution is becoming increasingly important. Empirical work reviewing more recent periods should account for this by treating mobile telephony as a potential substitute. Furthermore, fixed termination rates are low when compared to mobile termination rates, but not zero. Such termination rates may be above marginal costs, implying that termination costs and revenues need also to be considered in the profit maximising problem of the incumbent. As has been described earlier, this may imply waterbed effects where a decrease in wholesale prices may also lead to an increase in particular retail prices. In addition, an extended model could consider web-based VoIP as data becomes available as well as two-part tariffs. It should be noted, however, that the current limited availability of (quarterly) data points greatly restricts the complexity of an empirical model in this market. Any relaxation of the hypotheses described here would therefore not be implementable with the dataset used in this paper. While future work could reformulate the model and estimate it when more data points become available (or apply it to a market where more frequent data is available[40]), the present model represents a starting point, implementing the simplest possible specification with limited data.

[38] when each competitor has a price dependent hinterland of loyal customers unaccessible to the other operator.
[39] as well as wholesale line rental
[40] e.g. Austria, where monthly data is available

8. Bibliography

Agüero, A., de Silva, H., & Kang, J. (2011). Bottom of the Pyramid expenditure patterns on mobile services in selected emerging Asian countries. *Information Technologies & International Development, 7(3)*, pp-19.

Ahn, H., Lee, J., & Kim, Y. (2002). Estimation of a fixed-mobile substitution model in Korean voice telephony markets. Freie Universität Berlin, *Academic research, 3*.

Armstrong, M. (2002). The theory of access pricing and interconnection, in: M. Cave, S. Majumdar, and I. Vogelsang (eds.), Handbook of Telecommunications Economics.

Baker, J.B., & Bresnahan, T.F. (1985). The gains from merger or collusion in product-differentiated industries. *The Journal of Industrial Economics, 33*(4), 427-444.

Baker, J.B., & Bresnahan, T.F. (1988). Estimating the residual demand curve facing a single firm. *International Journal of Industrial Organization, 6*(3), 283-300.

Bresnahan, T. F. (1982). The oligopoly solution concept is identified. *Economics Letters, 10(1)*, 87-92.

Bresnahan, T. F. (1989). Empirical studies of industries with market power. *Handbook of industrial organization, 2*, 1011-1057.

BEREC. (2012). Report on the impact of fixed-mobile substitution on market definition. BoR (12)52

Briglauer, W., Schwarz, A., & Zulehner, C. (2011). Is fixed-mobile substitution strong enough to de-regulate fixed voice telephony? Evidence from the Austrian markets. *Journal of Regulatory Economics, 39(1)*, 50-67

Briglauer, W., Götz, G., & Schwarz, A. (2010). Can a margin squeeze indicate the need for deregulation? The case of fixed network voice telephony markets. *Telecommunications Policy, 34(10)*, 551-561.

Calzada, J., & Valletti, T. M. (2008). Network Competition and Entry Deterrence. *The Economic Journal*, 118(531), 1223-1244.

Davis, P., & Garcés, E. (2009). Quantitative techniques for competition and antitrust analysis. Princeton University Press.

Ecorys. (2013). Future electronic communications markets subject to ex-ante regulation. Study for the European Commission. Retrieved from http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=3148

Enders, W. (1995). Applied Econometric time series: Wiley.

Finnish Communications Regulatory Authority/PÖYRY. (2009). Mobile phone service prices 2009 – international comparison. 3/2009.

Forchheimer, K. (1908). Theoretisches zum unvollständigen Monopole. *(Schmollers) Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft, 32*, 1–12

Genakos, C., & Valletti, T. (2011). Testing the "waterbed" effect in mobile telephony. *Journal of the European Economic Association, 9(6)*, 1114-1142.

Gans, J. S., & King, S. P. (2005). Competitive neutrality in access pricing. *Australian Economic Review, 38(2)*, 128-136.

Gassner, K. (1998). An estimation of UK telephone access demand using Pseudo-Panel data. *Utilities Policy, 7(3)*, 143-154.

Growitsch, C., Marcus, J., & Wernick, C. (2010). The effects of lower mobile termination rates (MTRs) on retail price and demand. *Communications and Strategies, (80)*, 119-140.

Granger, C.W., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics, 2*(2), 111-120.

Gyimah-Brempong, K., & Karikari, J. A. (2007). Telephone Demand and Economic Growth in Africa. OECD Economic Studies, 38(1).

Inderst, R., & Peitz, M. (2012). Network investment, access and competition. *Telecommunications Policy, 36(5)*, 407-418.

Johnston, J., & DiNardo, J. (1997). Econometric Methods 4th Edition McGraw-Hill: New York.

Hjalmarsson, E., & Österholm, P. (2007). Testing for cointegration using the Johansen methodology when variables are near-integrated: International Monetary Fund.

Kahai, S.K., Kaserman, D.L., & Mayo, J.W. (1996). Is the Dominant Firm Dominant-An Empirical Analysis of AT&T's Market Power. *Journal of Law and Economics, 39*(2), 499-517.

Laffont, J. J., & Tirole, J. (2001). Competition in telecommunications. MIT press.

Lerner, A. P. (1934). The concept of monopoly and the measurement of monopoly power. *The Review of Economic Studies, 1(3)*, 157-175.

McCloughan, P., & Lyons, S. (2006). Accounting for ARPU: New evidence from international panel data. *Telecommunications Policy, 30(10)*, 521-532.
Nardotto, M., Valletti, T. M., & Verboven, F. (2012). Unbundling the incumbent: Evidence from UK broadband. Centre for Economic Policy Research.
Olsson, O. (2011). Estimating the demand and market power of a firm in sawn wood markets. *Aalto University Working Paper 1194*.
Qu, F. (2007). Empirical Assessment of Market Power in the Alberta Wholesale Electricity Market. *Energy Studies Review, 15(1)*, 4
Scheffman, D. T., & Spiller, P. T. (1987). Geographic market definition under the US Department of Justice Merger Guidelines. *Journal of Law and Economics, 30*.
Scherer, F.M., & Ross, D. (1990). Industrial market structure and economic performance. Houghton Mifflin.
Suslow, V. Y. (1986). Estimating monopoly behavior with competitive recycling: an application to Alcoa. *The RAND Journal of Economics*, 389-403.
Vickers, J. (2005). Abuse of Market Power. *The Economic Journal, 115(504)*, 244-261.
Vogelsang, I. (2003). Price regulation of access to telecommunications networks. *Journal of Economic Literature, 41(3)*, 830-862.
Vogelsang, I. (2010). The relationship between mobile and fixed-line communications: A survey. *Information Economics and Policy, 22(1)*, 4-17.
Ward, M. R., & Woroch, G. A. (2010). The effect of prices on fixed and mobile telephone penetration: Using price subsidies as natural experiments. *Information Economics and Policy, 22(1)*, 18-32.
Worcester, D.A. (1957). Why" dominant firms" decline. *The Journal of Political Economy, 65*(4), 338-346.
Yule, G.U. (1926). Why do we sometimes get nonsense-correlations between Time-Series? *Journal of the royal statistical society, 89*(1), 1-63.

9. Technical Annex

9.1. Estimation of reference model

In this technical annex, the detailed tests leading to the stable and interpretable model reported in the paper are presented. The Stata programme code to reproduce these results is reported in footnotes.

Second stage estimation (reference model)
For convenience the basic model equation (6) representing residual demand can be restated here.

$$q_D = \phi_0 + \phi_1 \hat{p} + w_F' \phi_F + w_C' \phi_C + x' \phi_B + \varepsilon$$

In this model no intertemporal effects are considered, i.e. changes of variables only affect the dependent variable of the same period. The second stage 2SLS estimations (robust standard errors) for residual demand using instrumented prices fitted in the first stage (see following section) are reported in Table 3. The estimation represents a regression of dominant firm quantities (q_D) on instrumented prices ($\hat{p}$) and other variables[41].

q_D	Variable	Coefficient estimate	Std. err.	P>| t |
$\hat{\phi}_1$	Price	-0.663	0.364	0.068
$\hat{\phi}_{B1}$	GDP per capita	-0.333	0.726	0.646
$\hat{\phi}_{B2}$	Number of active PSTN access lines	1.980	0.366	0.000
$\hat{\phi}_{B3}$	Seasonal dummies for quarters 2, 3 and 4	-0.032	0.016	0.042
$\hat{\phi}_{B4}$		-0.065	0.024	0.007
$\hat{\phi}_{B5}$		-0.021	0.013	0.112
$\hat{\phi}_{F1}$	Voice origination prices per minute	0.150	0.133	0.259
$\hat{\phi}_{F2}$	Number of wholesale ADSL lines	-0.220	0.043	0.000
$\hat{\phi}_{F3}$	Number of unbundled accesses	-0.001	0.001	0.548
$\hat{\phi}_{F3}$	Mobile termination rate (average)	0.116	0.071	0.104
$\hat{\phi}_{F4}$	Interest rates 30y Govenment bond	-0.050	0.052	0.338
$\hat{\phi}_{F5}$	Exchange rates EUR/CHF	-0.220	0.155	0.155
$\hat{\phi}_0$	_constant	4.390	4.102	0.285
R^2	0.9658			
Prob>F	0.000			
N	28			

Table 3 - Second stage estimation of residual demand (reference model)

Note that as preliminary regressions have indicated possible serial correlation of errors, robust standard errors are used (Huber-White-Sandwich estimator)[42].

[41] Stata command: ivregress 2sls lntradvoipallminscm lnyrealcapita lnpstn d2 d3 d4 lnscmregorr lnadslwhole lnullreal lnwavgmtrr lninterest30y lnfxratechfeur (lntradvoipallarpmallr=lnSCMstaff lnadslretail), first robust
[42] The estimations can be replicated (when omitting the robust command) by running a simple regression of the instruments on the instrumented variable, saving the fitted values and regressing the dependent variable of the second stage against the explanatory variables as well as the fitted variable. With robust standard error estimates though, results would slightly defer.

First stage estimation (reference model)

Similarly, for convenience basic model equation (5) representing instrumented prices can be restated here.

$$\hat{p} = \theta_0 + w_D'\theta_D + w_F'\theta_F + w_C'\theta_C + x'\theta_B + \varepsilon$$

The first stage 2SLS estimations (robust) for the regression of price on the instruments in the system are reported in Table 4.

p	Instruments	Coefficient estimate	Robust Std. err.	P>| t |
$\hat{\theta}_{B1}$	GDP per capita	*-1.450*	*0.620*	*0.035*
$\hat{\theta}_{B2}$	Number of active PSTN access lines	*2.063*	*1.280*	*0.129*
$\hat{\theta}_{B3}$	Seasonal dummies for quarters 2, 3 and 4	*0.033*	*0.020*	*0.133*
$\hat{\theta}_{B4}$		*0.052*	*0.019*	*0.015*
$\hat{\theta}_{B5}$		*0.005*	*0.033*	*0.884*
$\hat{\theta}_{F1}$	Voice origination prices per minute	*0.381*	*0.189*	*0.063*
$\hat{\theta}_{F2}$	Number of wholesale ADSL lines	*-0.363*	*0.231*	*0.139*
$\hat{\theta}_{F3}$	Number of unbundled accesses	*-0.003*	*0.003*	*0.353*
$\hat{\theta}_{F4}$	Mobile termination rate (average)	*0.111*	*0.167*	*0.517*
$\hat{\theta}_{F5}$	Interest rates 30y Govenment bond	*0.040*	*0.09*	*0.664*
$\hat{\theta}_{F6}$	Exchange rates EUR/CHF	*-0.261*	*0.257*	*0.327*
$\hat{\theta}_{D1}$	Number of staff working for Swisscom	*-0.225*	*0.185*	*0.243*
$\hat{\theta}_{D2}$	Number of retail ADSL lines active for Swisscom	*0.331*	*0.285*	*0.266*
$\hat{\theta}_0$	_constant	*-21.6948*	*10.074*	*0.049*
R^2	0.9298			
Prob>F	0.000			
N	28			

Table 4 - First stage estimation of residual demand (reference model)

What is of importance in this regression that there are significant instruments. This is the case.

9.2. Identification of econometric problems

Time series models are particularly complex to interpret as various time related problems can negatively affect the stability of the model and the extent to which it can be interpreted. Most importantly, two potential problems will need to be addressed: serial correlation of errors and non-stationarity of variables. This technical chapter analyses these problems and presents the solution adopted in the form of an ARDL model implying that the baseline equations are added a one period lagged dependent and independent variables.

Similar papers to this have often not proposed extended econometric analysis to check for the stability of results. In what follows it will be tried to address these problems in a structured way. Econometric modelling often implies a necessity for compromise as what desired by theory cannot be estimated efficiently. It will be shown that even though the sample size is limited, stable and meaningful results can be obtained.

When analysing the time series of the strategic variables graphically, it appears that there is a sufficient degree of variation in the locations of p and q_D to explain a demand and supply relationship (Figure 9).

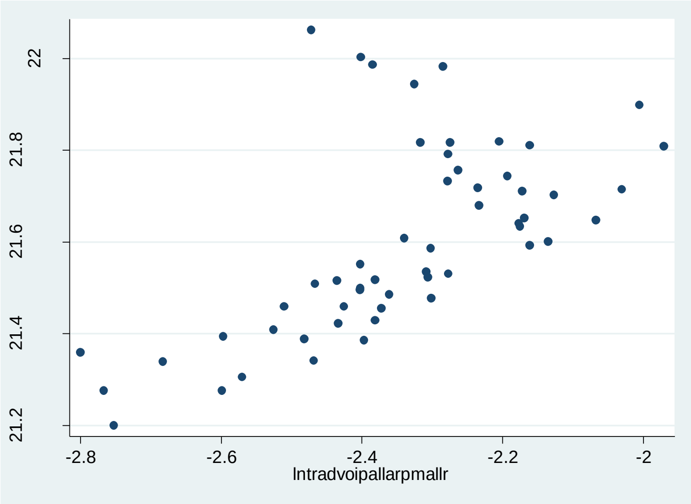
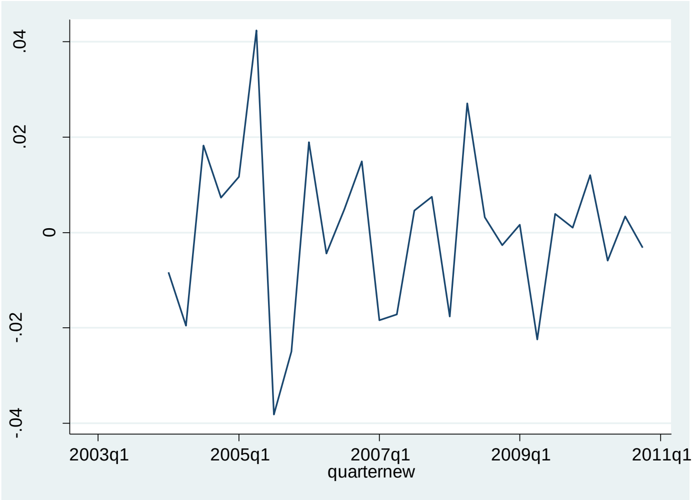


Figure 9 – Residual demand and prices in values[43]

The above graph shows the observed market outcomes (price is plotted on the horizontal axis and residual demand on the vertical). In the two following sections a general analysis of the time series under consideration is conducted and tests for the presence of non-stationarity and co-integration as well as serial correlation of errors are performed. Then, the econometric issues are analysed in detail and solution concepts are elaborated.

9.3. Test for serial correlation of errors

In case of serial correlation of errors, the errors tend to depend on their own lagged values. In the following it will be verified whether this is the case in the present regressions. The serial correlation of the second stage residuals is reported in Figure 10.



[43] twoway (line lntradvoipallminscm quarternew)(line lntradvoipallminoth quarternew)

Figure 10 – Graph of residuals over the period under consideration

From a graphical analysis there could be serial correlation of residuals, but the effect does not seem to be strong. A Portmanteau test for white noise confirms serial correlation of errors in the first stage regression (Table 5).

Portmanteau test for white noise [44]	
Portmanteau (Q) statistic =	25.1171
Prob > chi2(12) =	0.0143

Table 5 – Portmanteu test for serial correlation (first stage equation)

The test suggests rejecting the null hypothesis of no autocorrelation of errors (at a 5% critical value). There seems therefore clearly to be a problem of serial correlation.

9.4. Test for non-stationarity

It is important to know whether the series under consideration are stationary, i.e. have constant mean, variance and covariance over time or not, as this changes the framework of analysis required. The development over time of the dominant firm and fringe output is illustrated in Figure 11 (incumbent in blue, fringe in red), while the development in first difference is reported in Figure 12. The development of price levels of both Swisscom and competitor prices is represented in Figure 14 (incumbent in blue, fringe in red). Finally, Figure 13 reports the development of market shares of Swisscom.

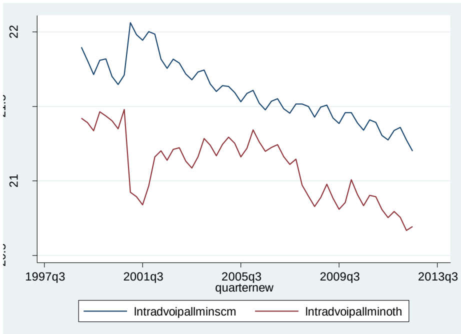


Figure 11 – Dominant firm and fringe outputs (national voice traffic minutes) [45]

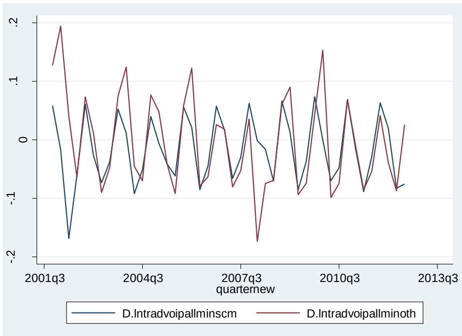


Figure 12 - Dominant firm and fringe outputs (national voice traffic minutes), first differences

[44] Stata command: wntestq res5 (where residuals are predicted after second stage using the true values of the endogenous variable.

[45] twoway (line d.lntradvoipallminscm quarternew)(line d.lntradvoipallminoth quarternew) if t>16

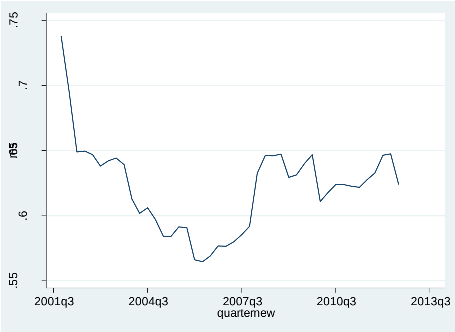


Figure 13 – Dominant firm market share on voice traffic minutes

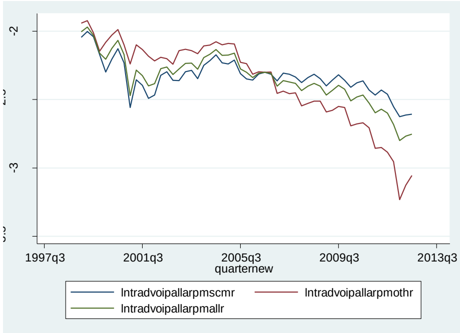


Figure 14 – Average prices (revenues) per minute

Figure 11 shows a general decline of the fixed telephony market which concerns both the incumbent as well as competitors. Overall, the output series are unlikely to be stationary (mean is decreasing over time). When looking at first differences it can be seen that there seem to be no changes in mean or variance (of first differences) over time. Figure 13 reports Swisscom's market share in detail (here for fixed national outgoing traffic only) which is in a declining market relatively stable over time as shown in the introduction on aggregate data. Finally, the analysis of prices from Figure 14 seem to suggest a decreasing trend and that except for the few last quarters under observation it may be reasonable to consider a single market price as prices move relatively closely together and it can even be seen that from around mid-period (2006) the sign of price differences between operators change and Swisscom perceives slightly higher per minute prices. As for also quantity, prices seem to decline over time not suggesting a stable mean and therefore unlikely to be stationary.

When looking at correlograms to identify how strongly variables correlate with their own "past" (autocorrelation) it can be seen in the following figures that all important strategic variables are significantly correlated with their own first lag. There seems to be, therefore, some kind of feedback effect in the market which has not been taken into account in the baseline model[46].

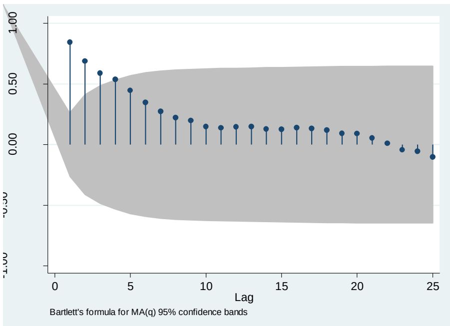


Figure 15 – Price AC

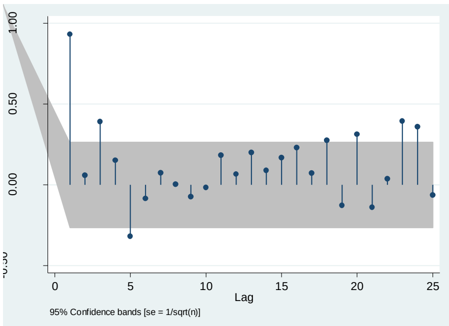
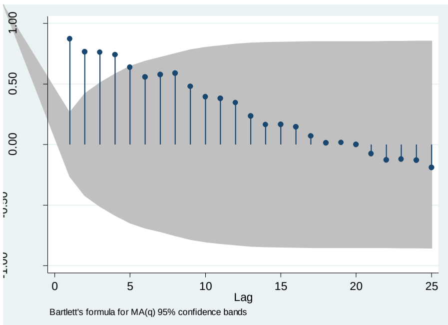
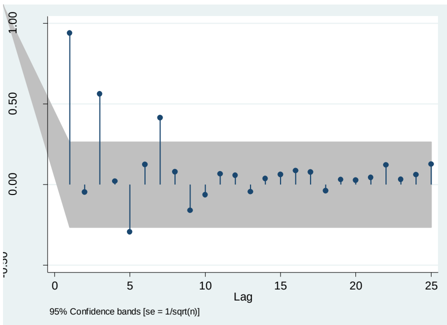


Figure 16 - Price PAC

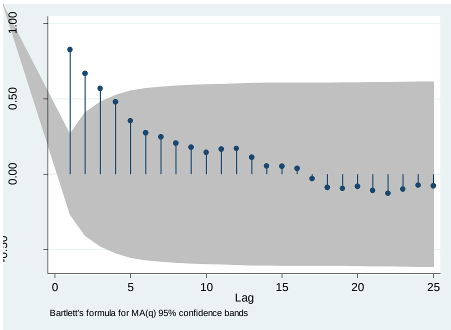




[46] AC=Autocorrelation, PAC=Partial autocorrelation

Figure 17 – Dominant firm output AC

Figure 18 - Dominant firm output PAC



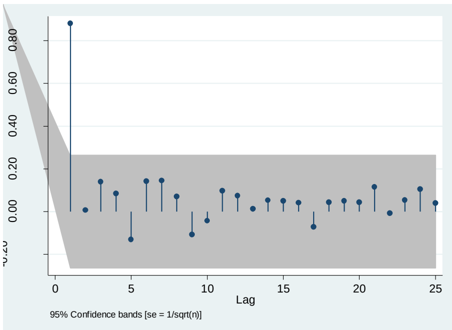


Figure 19 – Fringe output AC

Figure 20 – Fringe output PAC

While autocorrelations show that the correlation with current periods tends to decrease smoothly over time for all strategic variables considered, the partial autocorrelations (controlling for every single lag) show that the most important lags in explaining current levels is t-1, i.e. the preceding quarter for all variables analysed. Most importantly, Figure 17 and Figure 18 show that Swisscom's output strongly depends on its output in t-1 (it also depends relatively strongly on values in t-3). A similar effect applies to fringe output as can be seen in Figure 19 and Figure 20. This effect is also observed in Figure 15 and Figure 16 where prices seems to importantly depend on own values in the preceding period suggesting the same problem for the first stage regression. In general this analysis suggests that there is some kind of intertemporal effect which might need to be taken into account in the model.

Standard tests for stationarity include simple Dickey Fuller tests (no difference lags included). The Dickey Fuller test results are not reported here in detail. They suggest, however, that all variables used in the regression were, over the relevant time horizon of the regression, non- stationary, i.e. integrated of a higher order than zero. An exception is the variable "number of retail broadband subscribers of Swisscom" which is stationary as well as the constant term and the dummies). From all variables under examination in the period under review (4Q2003 to 3Q2012) all variables with an integration order higher than zero were, however, stationary in their differences and therefore integrated of order one. When performing an augmented Dickey Fuller test taking into account a series of lags these results are generally confirmed. The results overall indicate that there is a stationarity problem which has to be addressed.

9.5. Addressing serial correlation of errors

Under serial correlation of errors, the usual OLS estimators, although linear, unbiased and asymptotically normally distributed are no longer minimum variance among all linear unbiased estimators. As a result, the coefficient estimates remain valid but the usual t, F, and χ^2 statistics may be less efficient (higher probability of rejecting the true hypotheses and accepting wrong hypothesis). Serial correlation of errors may be caused among other things by a specification bias, excluding relevant variables from the regression (including lagged variables), non-stationarity of the series or incorrect functional form. There are different ways to address the problem of serial correlation. In the present model, the following approaches to structurally improve the model are of interest.

i) Using log-log specification

In some cases, the introduction of log-log specifications can improve results (as the procedure tends to shrink outliers) and lead to less serial correlation[47]. As logs are already used no further improvement is possible.

[47] Baker and Bresnahan (1988)

ii) Using first differences

The variables could all be first differenced. Such a transformation is, however, not compatible with the outlined model as coefficients of the level variables represent elasticities (e.g. residual demand elasticity). Differentiating these values would not allow the necessary inferences. For this reason this otherwise popular method is discarded.

iii) Adding new (lagged) variables

New variables, especially lagged variables could be introduced in order to effectively remove serial correlation of errors. Given the possibly dynamic nature of the model this could be a solution in one or both stage regressions[48]. Such transformations would be unlikely to change the model qualitatively and would only take into account the transmission effects of the function across time (in the baseline specification it is supposed that changes in t only have impacts in t). When analysing autocorrelation of the variables it is seen that dominant firm output seemed to be strongly correlated with its own first to third lag (the second lag is less correlated). Considering AC and PAC statistics the first lag in particular has been shown to be of particular importance.

Lagged variables (including the dependent variable) can be taken into account as exogenous (coming from t-1). Any introduction of a lagged dependent variable would mean that any other variable in the regression would have an “after effect” into the future. A change in any exogenous variable would then not only have a direct effect on the output demanded in t, but indirectly over the very change of the dependent variable in t also on, for example, the quantity demanded in t+1. A change in an independent variable would therefore result in an increase/decrease of the dependent variable to some extent also in t+1, t+2, etc. Long run multipliers measuring the long run marginal effects need in this case be calculated. Similarly it would also be possible to introduce lagged independent variables. This would then explicitly take into account the dynamics of the effect of a specific shock on the dependent variable. Considering the above, using lagged dependent and independent variables would be broadly compatible with the baseline model, simply introducing intertemporal effects. It has, however, to be considered that the low number of observation constrains the possibilities to include a large number of additional variables.

Finally, the model may simply continue to be estimated by OLS as coefficients are still unbiased, but the standard error may be corrected to be again minimum variance (e.g. using Newey West or Huber-Sandwich-Sandwich corrections to make the estimation of the standard error minimum variance, as used in the preceding estimations). The standard error is then called “robust”.

Overall, the option of introducing lagged variables seems a reasonable way to solve the problem. As such an option may, however, also have an influence on cointegration (see next section), possible adjustments to the model will be analysed subsequently.

9.6. Addressing non-stationarity

As described in Granger and Newbold (1974) stationarity (integration of order zero) is one of the underlying assumptions of a linear regression. In case of non-stationarity, when for example two otherwise independent variable grow over time (i.e. the mean is not constant), a significant statistical relation between the two can be found in a standard OLS regression when none exists. This not because these variables would have something in common but simply because both are growing (and this growth may be driven by outside variables). Yule (1926) has shown that spurious correlation would even persist in non-stationary series which are long. This effect could be controlled when introducing trends into the regressions. However, the theoretical model should be able to explain growth over time

---

[48] Olsson (2011)

rather than take it as exogenous and eliminate it. Introducing a trend would in the present case therefore not be appropriate.

If some variables in the regression are integrated of order one, then the usual statistical results may or may not be valid. Only in the case when regressors are integrated of order one (I(1)) and also "co-integrated", the regression can be reconstructed in a way that allows for valid inference without producing spurious results. Cointegration occurs when the long run stochastic trends in two processes are the same so they cancel. Variables then have a "common trend". In other words there must be a linear combination of the variables which is stationary (I(0)). Formally,

$$\beta_1 X_{1t} + \beta_2 X_{2t} + \cdots = 0$$

Where $\beta \;\; = (\beta_1, \beta_2, \dots, \beta_n)$ is called the co-integrating vector. The equilibrium error is the difference in t to the common trend. The equilibrium error for a given co-integrating relationship is $e_t = \beta \;\; x_t$. If x_t has n components there may be as many as n-1 linearly independent co-integrating vectors. When *x* contains only two variables, only one co-integrating relationship is possible. A matrix B may then include all co-integrating vectors. The number of co-integrating vectors of x is then called the cointegration rank of *x* (Enders, 1995). Intuitively, in the case of two of more co-integrated variables a shock on one variable would indeed have a relevant long term effect on the other variable in which case a regression may again be valid due to a common stochastic trend. There may therefore be departures from the long run equilibrium trend, but such a trend exists and the values of the variables tend to it over time. If there is not at least one such common trend between I(1) variables regressions using non-stationary variables are spurious.

An Engle-Granger test can be performed to check for cointegration of variables. In particular, it checks whether the residuals of the regressions are stationary or not (regressing the first difference of the residuals in t on the lagged level of residuals in t-1)[49]. In case of stationarity, the variables are co-integrated. When performing the test (Table 6 and Table 7), the hypothesis of a unit root in the errors in the baseline model in the first as well as in the second stage cannot be rejected[50].

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	**-5.221**	-6.846	-5.945	**-5.512**

Critical values from MacKinnon (1990, 2010)

Table 6 - Engle-Granger test for cointegration for the first stage regression[51]

	Test Statistic	1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	**-5.308**	-7.955	-6.857	**-6.340**

Critical values from MacKinnon (1990, 2010)

[49] Methodologically this test differs only slightly from the Dickey Fuller tests performed on the variables to check for their stationarity (using McKinnon Criterion).

[50] Note that it is formally incorrect using the I(0) variable which is included in our model (number of ADSL retail lines of the incumbent) in this test (or even in the model). However, the test results differ only slightly when excluding the variable. At a 5% critical value the hypothesis of non stationarity of the errors (Z(t)= -5.326)) could not be rejected. Ideally regressions should only be run using I(1) variables which are co-integrated with each other (full rank according to Johansen). In practice this is, however, rarely the case and regressions are also run without full rank and with few I(0) variables.

[51] egranger lntradvoipallarpmallr lnadslwhole lnullreal lnwavgmtrr lninterest30y lnfxratechfeur lnSCMstaff lnadslretail

Table 7 - Engle-Granger test for cointegration for the second stage regression[52]

The errors are therefore non-stationary and it can be suspected that neither the first nor the second stage equations in the baseline model are co-integrated. The results suggest therefore that the baseline estimates are the result of a spurious regression and are not properly interpretable. When interpreting such spurious regressions (such as the reference estimates in this paper), it can be considered that the estimation bias is always in the direction of rejecting a true null hypothesis (Granger and Newbold (1974)). Therefore, inferences leading to accepting the null hypotheses should usually be correct. In the case of the t test on coefficients (H0 is that the coefficient equals zero) this would imply that a connection between variables could be found which are in reality non existing. Instead, decisions to not reject the hypothesis that the coefficient is 0 are likely to be correct.

Finally, from a technical point of view regressions are meaningful when sufficient cointegration can be supposed – this even when including formally not integrated (or near integrated) variables (Hjalmarsson & Österholm, 2007). The following sections show how the problem of non-stationarity and absence of cointegration can be solved.

9.7. Addressing non-stationarity and lacking cointegration with an ARDL model

In most cases, non-stationary variables are difference stationary I(1), implying that the integration order of the differences is zero. This is the case also with the variables used in this paper. This automatically means that when using first differences instead of levels the non-stationarity problems of the regression disappear. Taking first differences is therefore a popular solution to the stationarity problem. As already described in the section on serial correlation of errors, the problem of this approach in the present paper is that information on levels are lost. While with a level estimation the estimated coefficients represent elasticities, in a first difference model coefficients would only indicate acceleration of such coefficients over time. Even in steady state, there would be no further information. Such an approach would, therefore, not be useful to determine residual demand elasticity and market power and is discarded.

This paper instead uses an Auto-Regressive Distributed Lags model to address the non-stationarity problem. Similarly to the autocorrelation issue the non-stationarity problem may be overcome by introducing lags of the dependent and independent variables, as long as there is cointegration. A simple first order ARDL model may represent a convenient bridge between a purely theoretical level model without lags and a more pragmatic data analysis considering intertemporal feedback effects. With dependent variable y and independent variables z, an ARDL (1,1) model, i.e. including one lag of the dependent variable and one lag of the independent variable, can be used (see Equation (9)).

$$y_t = c + \alpha_1 y_{t-1} + \beta_0 z_t + \beta_1 z_{t-1} + \epsilon_t \qquad (9)$$

In (9) $\beta_0$is called the “impact multiplier” (i.e. the immediate same period effect of the explanatory on the dependent variable). The other coefficients are dynamic multipliers representing how the system will adjust to the shock over time. The main objective of this paper is to identify the long term effects of changes in explanatory variables. The steady state is reached when the following equations hold.

$$y_t = y_{t-1} = y^*; z_t = z_{t-1} = z^* \; and \; \epsilon_t = 0$$

Then, the steady state is represented by the following equation:

[52] egranger lntradvoipallminscm xboriginal lnyrealcapita lnpstn lnscmregorr lnadslwhole lnullreal lnwavgmtrr lninterest30y lnfxratechfeur

$$y^* = \frac{c}{1-\alpha_1} + \frac{\beta_0 + \beta_1}{1-\alpha_1} z^*$$

In an ARDL(1,1) model a permanent change in *z* therefore affects *y* over all future periods, increasing it permanently. This overall impact, the "long run multiplier" is represented by (10).

$$\frac{\partial y^*}{\partial z^*} = \frac{\beta_0+\beta_1}{1-\alpha_1} \qquad (10)$$

Equation (10) can be estimated when the coefficients of the ARDL regression are known.

A possible problem when using ARDL for the model considered here could be that if the lagged endogenous variable appears on the right-hand side of the regression equation (as in (9)) and the disturbances continue to be autocorrelated, then the lagged endogenous variable will automatically be correlated with the disturbance term and thus become endogenous leading to biased and even inconsistent results. It the following sections it is, however, shown that the introduction of ARDL not only leads to sufficient cointegration but that in such a model errors are also not serially correlated anymore, meaning that this problem is inexistent.

Adaption of an ARDL Model

Given the nature of the theoretical model to be estimated (two equations, each with a series of explanatory variables to be estimated in two stages), it will first need to be analysed what type of ARDL model is most suitable. Then, it needs to be assessed whether introducing ARDL for both equations makes the first stage regression cointegrated and whether, therefore, a valid instrumented variable can be obtained. If this is the case, it also needs to be verified whether the second stage regression remains valid. Finally, it has to assessed whether there is still a problem of serial correlation of errors.

Choice of type of ARDL model

A first check of the relevant criteria to choose the number of lags to include for optimal cointegration (Table 8) suggests adding two to four time lags in the model (of the second stage equation using standard fitted values). This is not surprising as often in time series such tests suggest to introduce the maximum number of lags.

Selection-order criteria
Sample: 2005q1 - 2010q4 (N=24)

Lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	*369.533*				*4.60E-26*	*-29.9611*	*-29.8309*	*-29.4702*
1	*609.76*	*480.45*	*100*	*0.000*	*8.10E-31*	*-41.6466*	*-40.2142*	*-36.2472*
2	*3850.12*	*6480.7*	*100*	*0.000*	*1.e-140**	*-303.343*	*-300.609*	*-293.035*
3	*7335.47*	*6970.7*	*100*	*0.000*		*-591.289*	*-588.164*	*-579.508*
4	*7578.64*	*486.35**	*100*	*0.000*		*-611.553**	*-608.428**	*-599.773**

Endogenous: lntradvoipallminscm xboriginal lnyrealcapita lnpstn
lnscmregorr lnadslwhole lnullreal lnwavgmtrr lninterest30y lnfxratechfeur
Exogenous: _cons

Table 8 – Different criteria values to choose the number of lags to be considered[53]

Given the limited amount of data, the only *ARDL(k,k)* model that can be tested though is *k=1*, as including more lags in the model would make estimation unfeasible. Introducing only one time lag may not be optimal in resolving the problem of lacking cointegration, but there is a possibility that the problem is resolved. ARDL(1,1) regressions for the whole system are therefore run. In the following, it is shown that this model probably provides for sufficient cointegration.

Testing the first stage ARDL (1,1) regression for cointegration

The first stage equation, from (5) can be restated in ARDL(1,1) form in equation (11).

$$p_t = \hat{\theta}_0 + \hat{\theta}_1 p_{t-1} + w'_{D,t}\widehat{\theta_{D,t}} + w'_{D,t-1}\widehat{\theta_{D,t-1}} + w'_{F,t}\widehat{\theta_{F,t}} + w'_{F,t-1}\widehat{\theta_{F,t-1}} + w'_{C,t}\widehat{\theta_{C,t}} + w'_{C,t-1}\widehat{\theta_{C,t-1}} + x_t{}'\widehat{\theta_{B,t}} + x_{t-1}{}'\widehat{\theta_{B,t-1}} + \epsilon_t \quad (11)$$

The Engle-Granger test used in the preceding section is unfortunately limited to testing regressions with a low number of variables included. A simple test for cointegration as in the preceding section is therefore not possible here[54]. The alternative Johansen test on the regression (Table 9) shows, however, that there are six cointegration relations in the first stage regression, implying a relatively large amount cointegration relationships (with respect to the eleven variables included) providing evidence that spurious results may be avoided. This is particularly true with small samples. The Johansen test indicates, that the cointegration rank is relatively high. The results therefore show that a model taking into account one lag of all variables provides for sufficient cointegration and estimation of the instrumented variable should be valid[55].

Lags = 1
Trend: constant
Sample: 2004q2 - 2010q4, N =27

maximum rank	parms	LL	eigenvalue	trace statistic	critical value
0	*11*	*532.1482*		*544.8985*	*277.71*
1	*32*	*608.7789*	*0.99657*	*391.6371*	*233.13*
2	*51*	*663.0072*	*0.98199*	*283.1806*	*192.89*
3	*68*	*699.7184*	*0.93408*	*209.7582*	*156*
4	*83*	*726.9439*	*0.86691*	*155.3071*	*124.24*
5	*96*	*752.8149*	*0.85286*	*103.5651*	*94.15*
6	*107*	*772.2581*	*0.76313*	***64.6787****	*68.52*
7	*116*	*783.7952*	*0.57454*	*41.6046*	*47.21*
8	*123*	*794.3609*	*0.54281*	*20.4732*	*29.68*
9	*128*	*799.8902*	*0.33607*	*9.4145*	*15.41*
10	*131*	*803.885*	*0.25614*	*1.425*	*3.76*
11	*132*	*804.5975*	*0.05141*		

Table 9 – Johansen test for cointegration (first stage regression)[56]

Testing the second stage ARDL (1,1) regression for cointegration

The second stage equation from (6) can be restated in ARDL(1,1) form in (12):

[53] varsoc lntradvoipallminscm xboriginal lnyrealcapita lnpstn lnscmregorr lnadslwhole lnullreal lnwavgmtrr lninterest30y lnfxratechfeur
[54] Some authors also use simple Dickey Fuller tests to check for stationarity of the residuals.
[55] Although Johansen's methodology is typically used in a setting where all variables in the system are I(1), having few stationary variables in the system is theoretically not an important issue and Johansen (1995) states that there is little need to pre-test the variables in the system to establish their order of integration. If a single variable is I(0) instead of I(1), this will reveal itself through a co-integrating vector whose space is spanned by the only stationary variable in the model.
[56] Stata command: vecrank lntradvoipallarpmallr lnyrealcapita lnpstn lnscmregorr lnadslwhole lnullreal lnwavgmtrr lninterest30y lnfxratechfeur lnSCMstaff lnadslretail, lags(1)

$$q_{D,t} = \hat{\phi}_0 + \hat{\phi}_{1,t}\hat{p}_t + \hat{\phi}_{1,t-1}p_{t-1} + \hat{\phi}_{2,t-1}q_{t-1} + w'_{F,t}\hat{\phi}_{F,t} + w'_{F,t-1}\hat{\phi}_{F,t-1} + w'_{C,t}\hat{\phi}_{C,t} + w'_{C,t-1}\hat{\phi}_{C,t-1} + x_t{}'\hat{\phi}_{B,t} + x_{t-1}{}'\hat{\phi}_{B,t-1} + \epsilon_t \quad (12)$$

The Johansen test for this regression is reported in Table 9 and shows for the second stage that there are four co-integration relations (ten variables have been included).

Lags = 1
Trend: constant
Sample: 2004q2 - 2010q4, N =27

maximum rank	parms	LL	eigenvalue	trace statistic	critical value
0	*10*	*482.4242*		*384.8823*	*233.13*
1	*29*	*544.5308*	*0.98995*	*260.6692*	*192.89*
2	*46*	*579.4391*	*0.92466*	*190.8526*	*156*
3	*61*	*609.6892*	*0.89362*	*130.3523*	*124.24*
4	*74*	*630.6383*	*0.78813*	***88.4542****	*94.15*
5	*85*	*646.6574*	*0.69474*	*56.4159*	*68.52*
6	*94*	*656.4064*	*0.51429*	*36.9178*	*47.21*
7	*101*	*663.9514*	*0.42815*	*21.8279*	*29.68*
8	*106*	*669.2342*	*0.32383*	*11.2624*	*15.41*
9	*109*	*673.9483*	*0.29475*	*1.8341*	*3.76*
10	*110*	*674.8654*	*0.06567*		

Table 10 - Johansen test for cointegration (second stage regression) [57]

The results also show that in the second stage there is a consistent amount of co-integrating relationships and that the estimates of the second stage should also be valid.

Testing the ARDL (1,1) equations for serial correlation of errors

To conclude, it should be verified whether the ARDL (1,1) regressions continue to suffer the same problems of serial correlation of errors as the baseline model. All one period lagged variables (t-1) newly included in the regression can be considered exogenous, both in case the current period variables are exogenous or endogenous. The lagged dependent variable in the first stage (price) can therefore also be considered as an exogenous instrument (“coming” from the past)[58].

When testing the first stage instrumented variable equation for serial correlation of errors, no such serial correlation is found at critical values of 5% (Table 11).

Portmanteau test for white noise [59]

Portmanteau (Q) statistic = 18.2777
Prob > chi2(11) = 0.0754

Table 11 - Portmanteu test for serial correlation (first stage ARDL equation)

Overall, the first stage ARDL (1,1) regression is therefore estimating instrumental variables with instruments that are non-stationary, but having a sufficient degree of cointegration. In addition, there is no serial correlation of errors. The fitted value estimates produced in the first stage are therefore based

[57] Staat command: vecrank lntradvoipallminscm lnyrealcapita lnpstn lnscmregorr lnadslwhole lnullreal lnwavgmtrr lninterest30y lnfxratechfeur xboriginal, lag(1)
[58] The lagged price variable is therefore represented by real and not fitted values.
[59] Stata command: wntestq ress (where residuals are predicted after first stage simple regression).

on coefficients of a valid OLS regression, unbiased and efficient. It should, however, be noted that the significance of the instruments is reduced with respect to the first stage baseline estimates (given the now large number of variables this is not surprising).

As a final test for stability, a test for serial correlation of errors of the second stage regression is performed (see Table 12).

Portmanteau test for white noise [60]	
Portmanteau (Q) statistic =	*13.2888*
Prob > chi2(11) =	*0.2749*

Table 12 - Portmanteu test for serial correlation (second stage ARDL equation)

The results indicate that not only in the first stage, but also in the second stage, serial correlation of errors is not an issue anymore when using the ARDL(1,1) specification. Nevertheless, standard errors continue to be estimated with a "robust" correction as before, as doing so may only further strengthen results.

The second stage ARDL (1,1) regression is therefore estimated with variables that are non-stationary, but having a sufficient degree of cointegration. In addition there is no serial correlation of errors.

Overall both the first and second stage ARDL (1,1) regressions are therefore valid. A full ARDL (1,1) model may consequently be used to correct the baseline model for the econometric problems identified.

[60] Stata command: wntestq ress2 (where residuals are predicted after second stage using the true values of the endogenous variable).